DOCUMENT RESUME

ED 036 659 AA 000 497

AUTHOR	Turner, Edward F., Jr.
TITLE	A Study of the Implications of Modern Technology for Small College Libraries.
INSTITUTION	Washington and Lee Univ., Lexington, Va.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-7-0910
PUB DATE	31 Mar 69
GRANT	OEG-1-7-070910-3706
NOTE	139p.
EDRS PRICE	EDRS Price MF-$0.75 HC-$7.05
DESCRIPTORS	Bibliographic Citations, *College Libraries, Computers, Information Storage, *Information Systems, *Library Networks, Library Services, Microforms, *Resources, *Technology

ABSTRACT

This project proposes to ascertain the extent to which the tools of modern technology can be applied to help solve the problems of small college libraries. For the purpose of this report, the small college library is one serving undergraduate instructional needs as distinguished from the university library, serving the interests of serious research. The first chapter of the report seeks to define clearly the relationship of library service to college instruction. This approach indicates the nature and magnitude of changes which have taken place in library service and demonstrates why the role of the library has become a central issue. The second chapter examines the nature of technology and its relation to library operation. Chapters 3 and 4 discuss the nature of library collections with a view to identifying some level of commonality of problems and collections, and to identify the advantages and disadvantages of major adaptations of technology. Chapters 5 and 6 look at special problems in the areas of bibliographic access and physical access. The report concludes with a summary which identifies areas requiring further investigation, and a series of recommendations. The primary recommendation of the report is the package library, built on the core of disciplines, and common to all colleges offering the same program. (Author/ON)

AA 000 497

Final Report

Project No. 7-0910
Grant No. OEG-1-7-070910-3706

A STUDY OF THE IMPLICATIONS OF MODERN TECHNOLOGY FOR SMALL COLLEGE LIBRARIES

Project Director: Edward F. Turner, Jr.
in collaboration with Stanley McElderry and William Kurth

Washington and Lee University
Lexington, Virginia

March 31, 1969

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

ED 036 659

Final Report

Project No. 7-0910
Grant No. OEG-1-7-070910-3706

A STUDY OF THE IMPLICATIONS OF MODERN TECHNOLOGY FOR SMALL COLLEGE LIBRARIES

Project Director: Edward F. Turner, Jr.
in collaboration with Stanley McElderry and William Kurth

Washington and Lee University
Lexington, Virginia

March 31, 1969

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

CONTENTS

Page
Preface v
Summary ix

Introduction 1

Chapter

I The Role of the Library in College Education 5
II Technology and the Library 17
III The Core Collection 23
IV Approaches in the Use of the Core Collection 43
V Bibliographic Access 63
VI Physical Access 75
VII Summary and Recommendations 93

Appendices

I Theoretical Basis for a Core Collection 101
II Citation Analysis of "Social Stratification" 119

Bibliography and References 127

Tables

I Colorado College and University Libraries and Duplication 28
II Washington, D.C. University Libraries and Duplication 30
III New England University Libraries and Duplication 30
IV Number of Books Reviewed Annually in Choice 36
V Individual Library Costs Compared to Center Operation 53
VI Variation in Cost per Volume According to Number of Copies Processed Simultaneously 55
VII Variety of Computer Components 72
VIII Introductory and Advanced Texts and Monographs Used in Citation Analysis 120
IX Sample List of Duplication in Citation Analysis 121
X Frequency of Citation in Sociology Sample 122
XI Complete Matrix of Duplicate Citations 123

Figures

1. Analysis of 28,000 Annually Produced U.S. Books 35
2. Graphic Models of Centralized Library Centers 46-47
3. Processing Materials Extracted from Machine Readable Bibliographic Record 71
4. Block Diagram of a Computer 79

Preface

As a project of this sort comes to an end, I suppose it is inevitable that one becomes aware, for the first time, how great is the discrepancy between wish and fulfillment. The questions are more numerous now than at the beginning. There is, of course, one solid accomplishment - the education of a physicist in some of the complexities of library science, an education which has perhaps left him a wiser and less-assured person. The reader must be judge of whether anything else was accomplished, but I will state briefly what was attempted.

Much has been written about revolutionary changes in conventional library procedures. Some writers speculate that libraries as we now know them are a thing of the past and will ultimately be replaced by computers and electronic transmission of information from a central storage point. Others visualize the Library of Congress stored in a shoe box by use of ultra-microform which can be purchased for a fraction of the amount now expended annually by college libraries. Governmental agencies and private foundations are devoting considerable effort to determine the optimum method for handling large bodies of information, and experts from a wide variety of disciplines outside the library profession are engaged in seeking solutions to methods of handling the flood of information which large research labraries must accumulate and service.

There appeared to be a need to assess these developments and proposals. It was felt that an adequate assessment required a clearer perspective of the requirements of the college library and the alternative solutions for meeting these requirements than was available to those college administrators, faculty members, and librarians who have the responsibility for making the decisions to act upon these developments. This report attempts to provide, in non-technical terms, the background material on which informed judgments can be based.

My attention was first drawn to the pressing problems of academic libraries by Dr. Fred C. Cole, several years ago. Dr. Cole, now president of the Council On Library Resources, was then president of Washington and Lee University. He had, even at that time, a long familiarity with the needs of research libraries, and with the various proposals and research efforts then in progress that were directed toward the alleviation of those needs. It was an act of considerable trust and faith for him to invest so much of his time and energy to the education of one who could bring so little to the enterprise in the hope, I suppose, that something in the training and approach of a scientist could ultimately contribute to some solution of the library problem. Whatever his reason, it was a compliment which I appreciated, and his generous support of all my activities since that time is gratefully acknowledged.

My earlier investigations had convinced me that scientists and other technical people working on library problems had not always had access to the kind of expertise which is necessary if the proposed solutions are to be practicable as well as feasible. It was a condition of this grant, at my request, that this project be undertaken in close working collaboration with a librarian of the highest professional qualifications. It was my good fortune to have Stanley McElderry with me in that capacity throughout almost the entire project. At the time he joined in this endeavor he was librarian at San Fernando State College and a consultant to the California State board that directed the libraries of the nineteen colleges of that state. In September, 1968, he was appointed Dean of the Library School at the University of Texas. Despite the pressures of his other jobs, he maintained an intense interest and devoted an enormous amount of his time and energy to this project.

About midway through the study, we were joined by Dr. William Kurth, associate director of libraries at Washington University. He, too, has been indefatigable in both interest and effort. Messrs. McElderry and Kurth have been listed as collaborators in this project in partial recognition of the debt I owe them. Whatever have been the deficiencies of this project or this report, however, the responsibility is mine.

Many others have been generous in their aid and support. A number of faculty members of Washington and Lee were consulted, often without compensation, so many in fact, that I will not trust my memory to list them all, bu the support of Professors Kimbrough and Thomas of the Sociology Department, of Professor Chaffin of the Speech Department, and of Dr. Maurice Leach, university librarian, deserves special mention.

I am indebted to the administration of Washington and Lee for the release of time from my teaching duties to engage in this research, and for their understanding and interest; and to various administrative and service offices, notably the treasurer's office and the Central Mimeograph office. The genuine and sympathetic interest in the project on the part of President Robert Huntley is greatly appreciated.

Mr. Carl Spaulding and Mr. Melville Ruggles, of the Council on Library Resources provided much help. The former was my chief authority on the technology of microfilm, and the latter supplied much encouragment and information in my studies on the package library.

Dr. Malcolm Sillars, now acting president of San Fernando College, closeted himself with the principals of this study for a week in a Washington hotel room, hammering out the final revision of this report. Whatever lucidity and clarity the report conveys is due him in large measure.

As usual, probably the hardest working people of our project were the two ladies who served as clerks, secretaries, typists, key-punch operators, and research assistants - Mrs. Marcia Putney and Mrs. Wanda Kirby. They carried out every request, reasonable or otherwise, with aplomb.

To the other staff members of the physics department, who must have wondered at times whether I was librarian or physicist, I extend my thanks, not only for their patience but for their advice and assistance. And to my wife and family for their forbearance and encouragement. Much of them has gone into this project too.

Summary

This project was initiated to ascertain the extent to which the tools of modern technology can be applied to help solve the problems of small college libraries. In the context of this report, the small college library is one serving undergraduate instructional needs as distinguished from the university library, serving the interests of serious research. The study focussed on the "core collection", a concept which suggests there is a basic set of library resources which any college library ought to have in order to perform its objectives effectively. Efforts were directed toward determining the validity of the "core" concept, assessing its effectiveness and its comparative economic advantage over existing systems, and the manner in which the newer technologies, especially the computer and microform, were more or less adaptable to it than to present methods.

The study has led to the conslusion that the nature of higher education has changed in such a way as to call academic libraries into being and to require extensive use of libraries in the instructional process. The changing nature of curricular and instructional methods, the increased student population, and the accelerating rate of publication have produced a complexity in library operations which makes it difficult for them to carry out their responsibilities by using the techniques currently employed.

The discipline of Information Science has emerged in response to this complexity and to various attempts to utilize machines to solve the problems of libraries. The initial efforts have been addressed to the needs of science and technology, where the need was greatest, and where the literature was most tractable to analysis. Academic Libraries, encompassing the information requirements of all users, are not so amenable to the methods developed up to now by Information Scientists. Technology is most effectively applied to the solutions of problems where coordination and mass production can be achieved. The feasibility of applying technology to most, if not all, library operations has been demonstrated. However, coordination requires the exestence of a central agency, a step not yet achieved.

The establishment of a consensus about the resources which it is appropriate for college libraries to have opens the door to mass production for the distribution of bibliographic and physical access to the literature. Current methods of defining resources are deficient because of lack of knowledge of the functional relationship between a discipline and its literature. The best foreseeable way of defining resources is by a consensus of subject experts, because disciplines are determined that way. Technology, by monitoring use, can be effectively applied to refine the consensus of subject matter experts.

The computer and the production of bibliographic data in machine-readable form offer the most immediate solution to the problem of bibliographic access. It is feasible now to employ computer technology in this way but not economically sound for individual college libraries to attempt it. When accomplished, however, computerized bibliographic information offers the added potential to automate cataloging and the physical processing of library materials as well as providing increased opportunities for sharing of resources.

The present media for storing indexes and documents impose a bulkiness that increases costs and limits accessiblity. Methods of sharing resources with other libraries have not been successful to date. Two methods for providing access to index and textual material are the computer and microform. The computer is not appropriate for storage of documents but is the ideal instrument for handling index information. For the immediate future, books will remain as the primary means of document store. Microform has use as a storage medium for little-used materials now, and has potential for more extensive use. Various attempts to physically combine index and document store are undersirable.

The most fruitful avenue for effectively utilizing technology in libraries is to exploit mass production techniques through coordination of effort. A two step process for doing this is suggested, culminating in the "package" library. Various research activities that will enhance the probability of bringing this about are suggested.

Introduction

This project was initiated to ascertain the extent to which the tools of modern technology can be applied to solve the problems of the small college or community college library. It was anticipated that the project would identify those problem areas in the organization and operation of the college library which would lend themselves to technological solution and to make recommendations for action.

A college is frequently defined by size of enrollment, but for purposes of this study a functional definition seemed to be more appropriate. The college library, in the context of this report, is one serving undergraduate instructional needs as opposed to the university library serving the broader interests of serious research.

Attention has been centered on college libraries because they represent the majority of academic libraries and they have received less attention than the large research libraries. As a group they are the most seriously deficient in meeting current standards of adequacy. Although the financial problems of these libraries are acute, their information needs are less complex than for other types of libraries and they are more tractable to analysis in assessing alternative solutions to their problems.

Exploration of the literature on library technology, visits to library projects employing advanced technology, and discussions with various equipment manufacturers, librarians and library researchers demonstrated that considerable experimental investigation is already underway. Little of the effort is directed toward the problems of the small academic library, however. A more precise definition of the needs of college libraries and identifcation of areas for the application of technology appeared to be a more fruitful avenue of approach for this study.

Interest in the problems of small college libraries arises from the apparent inability of these libraries to keep pace with increased demands for service, which are mounting faster than can be met by continued reliance upon traditional approaches.

The most obvious difficulty to the library is that of identifying and securing the significant portion of the flood of new information stemming from the accelerating pace of research activity. While access to increasing amounts of this information is undoubtedly essential, the paramount problem of undergraduate institutions is the identification and acquisition of that part of this new information which is most relevant to their purposes.

A second element in the increased demand for library services is the number of persons to be served. College enrollments have risen sharply in the past decade and continue to rise. There are not only more persons of college age in the total population, but a higher percentage of them are attending college. Changes in college enrollments are also reflected in the number of faculty depending on the library for services.

Third, in addition to these quantitative changes there have been profound changes in curricula and in academic requirements. The increased amount of knowledge, increased fragmentation of traditional disciplines, and the changes in vocational requirements have resulted in the introduction of new subject matter and changes in instructional methods. The tendency has been to increase academic standards and to place a larger responsibility on the student for acquisition of information.

If academic libraries are to meet these pressures, the acquisition rate must be increased to keep pace with the production of information; greater scope and depth of collections are required to match corresponding changes in the curriculum; and more copies of individual titles are required to serve larger enrollments. Even in undergraduate faculties, faculty members are increasingly involved in research activity and require library support in this endeavor.

Clearly, larger library budgets are required to develop and maintain collections; to hire the staff to process and service these collections; and to provide the additional space to house them. All of these needs arise at a time when costs of books, personnel, and space are rising sharply.

These pressures are felt especially in the smaller academic institutions which face corresponding cost increases in other areas of activity. The library needs a larger percentage of the budget to maintain or improve quality, at a time when other divisions of the college are also attempting to up-grade services to meet insistent requirements of larger enrollments in the face of rising costs. Without new sources of revenue in increased amounts the smaller academic institutions face relative deterioration in quality of service. However, since a large percentage of college libraries are already seriously deficient in resources, and because libraries now play a central role in the instructional process, some solution to library problems must be found despite the demands of other pressing needs of academic institutions.

The first chapter of the report seeks to define more clearly the relationship of library service to college instruction. The procedure followed has been to examine, in some detail, the history and trends in educational philosophy, curricula and instructional methods. This approach indicates the nature and magnitude of changes which have taken place and demonstrates why the role of the library has become a central issue.

The second chapter examines the nature of technology and its relation to library operation.

Chapters 3 and 4 discuss the nature of library collections with a view to identifying some level of commonality of problems and collections, and to identify the advantages and disadvantages of major adaptations of technology.

Chapters 5 and 6 look at special problems in the areas of bibliographic access and physical access.

The report concludes with a summary, the identification of areas requiring further investigation, and a series of recommendations.

Chapter 1

The Role of the Library in College Education

The importance of the library as a means of instruction is primarily a phenomenon of the last century. Initially, the library served little function. Although book collections existed in most, if not all, colleges from their founding, they were small in size and little used. "Starting with 300 volumes, Harvard had only 12,000 by the conclusion of the Revolutionary War", say Brubacher and Rudy. "But even at that its library was far larger than its nearest rival, William and Mary, with 3000 or Dartmouth with 1000, the smallest of the then college libraries." (10, p. 94)

The reason for this poverty in book collections is not only that books were scarce generally and had to be imported from abroad, but also that the nature of instruction required little access to libraries. The pattern of instruction, borrowed from the English college, persisted until well into the nineteenth century.

> The two most popular methods of instructing during class periods were the recitation and the lecture. Although more popular in the seventeenth and eighteenth centuries, the recitation method gradually yielded ground to the lecture method, especially in the nineteenth century. The heart of the recitation consisted in an exchange between the tutor and the student, the tutor citing and the student reciting. The citation was usually an assignment in a textbook, but might just as well be a previous lecture or scientific demonstration. In the recitation the student proved that he had learned his lesson, at least the portion for which he was called on in class. (10, p. 82)

These methods of instruction were applied to a narrow prescribed classical fare, borrowed from the English college and deemed adequate for the training of ministers and community leaders, which was the sole purpose of higher education at the time. Henderson observed that until the end of the nineteenth century

> the prevailing philosophy among college educators in the United States was to educate a selected

> group of students of high ability. This philosophy can be recognized as the European, or classical, tradition. Under this theory a limited group of young people would be educated at the university where they would be inculcated with the cultural heritage. They, then, would become the leaders of society - in government and the professions - thus assuring cultured and presumably wise leadership. The remainder of the youth needed little education, or at most, some vocational education for their work. (29, p. 6)

Limited funds, and the scarcity of books might justify perpetuation of these methods of instruction. However, despite the improved quality of students and faculty, new knowledge, greater supply of books, and a demand for other types of training, the classic 1828 Report of the Yale Faculty stoutly defended, on philosophical grounds, the limited curriculum and instructional methods initially imposed by economic constraints.

> 'The two great points to be gained in intellectual culture, are the discipline and the furniture of the mind; expanding its powers, and storing it with knowledge', they said. These points might be best gained by adherence to the ancient subjects, for these were the subjects most certain to discipline and most worthy to furnish a balanced mind. Mathematics shaped the mind as an instrument of reasoning. The classics helped to achieve balance by bending the mind toward taste. In answer to the question of why each student should not 'be allowed to select these branches of study which were most to his taste, which are best adapted to his particular talents, and which are most nearly connected with his intended profession', the Yale people had a ready answer. 'Our prescribed course contains those subjects only which ought to be understood... by everyone who aims at a thorough education'.
>
> . . .Even the use of textbooks and recitations was defended as being superior to an arrangement that would send students into libraries, there to read the sources and conflicting authorities: 'The diversity of statements in these, will furnish the student with an apology for want of exactness in his answers.' (49, pp. 132-134)

The Yale Report provided authority for many colleges to resist changes in programs and instructional methods. Departure from established traditions occurred in new institutions such as Union College in 1802, University of Pennsylvania in 1816, and Rensselaer Polytechnic Institute in 1824, these being more hospitable to modern languages, new knowledge in the sciences, and training in mercantile and mechanical occupations. However, the impetus of Jacksonian democracy, public acceptance of higher education, and public support of education at the state and national level were required in order to generate the variety of programs and to attract the requisite number of students needed to support them. These conditions likewise provided the economic base for the introduction of change in other elements of the academic milieu such as consideration of the student environment, new instructional methods, and improved facilities. But the emphasis of libraries was to wait still longer.

College libraries did not attain significant size or function until the latter part of the nineteenth century. In 1876 Harvard held 212,050 volumes, Yale 95,200 and no other library has as many as 50,000 volumes. (23, pp. 3-16) Brubacher and Rudy indicate that rules for the use of college libraries down to the time of the Civil War

> ...allowed the library to be open only certain days in the week and for only a very limited amount of time even then. In the eighteenth century the Dartmouth Library was open for an hour once every fortnight and only five students were admitted at a time. Freshmen were permitted to withdraw only one book at a time, sophomores and juniors two, and seniors three. Yet is is startling to realize that a century later the rules for the use of the Princeton Library were only slightly more liberal. The predominant recitation method made no demands on the library. Assignments were made in textbooks of which each student had his own. No one thought of the library as a place to study. Dusty, ill-lighted, and unheated, it would have been an uninviting place even if it had been open more frequently and regularly. (10, p. 95)

An indication of the role the library would eventually play in college instruction is provided by the activities of the Enlightment-oriented student literary societies. Rudolph says that the debating club or literary society constituted

> the first effective agency of intellect to make itself felt in the American college...
>
> The first debating clubs probably owed something of their origin to the general atmosphere of colonial political debate that surrounded their birth and something to the continued lively political interest of the times. In debating clubs students could face squarely the exciting issues of the day, issues that occupied the attention of their elders in the partisan press, at the village tavern, or around the cracker barrel of the country store. The tradition of the medieval disputation which had been declining as a commencement activity also found new lodgment in the debating societies, and the oratorical and declamatory exercises of the curriculum received new encouragement. But it was the Enlightment faith in intellect, the commitment to reason, that most accounted for the development of the literary societies as a characteristic expression of undergraduate life...
>
> In their libraries the literary societies made most clear the degree to which the purposes they served were alien to those of the college. Not only did the literary societies often outstrip the college libraries in number of volumes, but the wide range of subject matter allowed far greater opportunity for the play of intellect than did the narrow religious fare of the usual college library. In 1835 the literary societies at the University of North Carolina contained 6,000 volumes, the best collection in the state. By 1840 the literary societies at Bowdoin boasted between 5,000 and 6,000 volumes; at Brown 3,000; and at William and Mary 10,000. In every case, these libraries were superior to the college libraries. Works of fiction, history, politics, and science were available to students because their literary societies purchased them...(49, pp. 136-144)

The literary societies demonstrated the utility of libraries but other, more pervasive, factors extended the concept. The rationale which made the college library a necessity, opened the flood gates of knowledge, and gave academic respectability to a wide variety of new subjects, was found in "the new psychology", and President Eliot at Harvard became one of the most effective advocates of change. In his inaugural address he proposed acceptance of the revolutionary elective system.

> The elective system fosters scholarship, because it gives free play to natural preferences and inborn aptitudes, makes possible enthusiasm for a chosen work, relieves the professor and the ardent...(student) of the presence of a body of students who are compelled to an unwelcome task, and enlarges instruction by substituting many and various lessons given to small lively classes, for a few lessons many times repeated to different sections of a numerous class. (49, pp. 292-293)

The elective system anticipated the psychology of individual differences, later demonstrated in experimental psychology, and repudiated the notion of transfer of training. Although Eliot's rationale "rested on a combination of desire, necessity, principle, and preference," (49, p. 293) it was in tune with the democratic spirit of the times which was opposed to aristocratic concepts of education and believed in the equality of all men with respect to recognition and opportunity to pursue the occupation of their choice. The elective principle was bitterly contested by some as detrimental to higher education, but in the end was accepted in one form or another by the majority of institutions. The impact of the adoption of the elective principle was the extension of the scope and depth of the curriculum, addition of new faculty, and augmentation of laboratory and library facilities. Increases in enrollment offset the cost of this expansion in the larger institutions but unfortunately many of the smaller colleges were forced to decide the issue more on financial than philosophical grounds.

Reforms in the curriculum touched off by the elective principle were matched by similar innovations in instructional methods. Brubacher and Rudy claim that:

> this revision represented a shift in the educational center of gravity from subject matter to the student. While older methods took the intrinsic values of subject matter for granted and therefore concentrated on transferring them to the student, newer methods were more inclined to take hold of the interests of students as the handle to teaching the values inherent in the curriculum. Older methods, for the

> most part, not only disregarded the interest of the student but even counted it a point in their favor that they disciplined the student to apply himself to subject matter without or even against the inclination of interest. (49, p. 271)

With students not forced into a specified curriculum and such secondary motivations as grades, exclusion from campus activities, dean's lists, and honors being found unsatisfactory

> The goal of reformers was to promote studies to the position of primary stimulant, where the student would see them as the direct means to satisfying his personal longings and ambitions. Should such an outcome occur, the objectives of college education would play a much more important part than ordinarily realized. They would not only give direction to the students' educational effort but they would also supply it with motive power. But to be sure that that power reached its destination, the objectives had to be the students' own. (49, p. 273)

A variety of approaches resulted from the efforts to center instruction on the interests of the student. The problem method became the most popular approach after the turn of the century. (49, p. 273) John Dewey was the most influential in promoting this method. In his view

> The point of departure...was to choose some perplexity in which the student had a sense of involvement. Such perplexities might run all the way from large scale problems taken from public life to the less pretentious ones of private life. Thus, intrigued by the tensions between East and West, a student might want to study some aspect of Russia as it impinged on the western world. Or he might want to probe deeply into the philosophical or psychological ramifications of some moral strain agitating him personally. Once past the point of departure, then followed the familiar steps in the problem method of gathering data, forming hypotheses, and testing them so far as academic limits permitted. (49, pp. 273-274)

There were also efforts made to provide a meaningful interrelationship between learning and activity whereby a student could apply knowledge as it was learned.

> "In the process of learning there should be present, in some sense or other, a subordinate activity of application," according to Alfred North Whitehead. "In fact, the applications are part of the knowledge, for the very meaning of the things known is wrapped up in their relationships beyond themselves. Thus unapplied knowledge is knowledge shorn of its meaning." (49, p. 274)

Perhaps most important to our concern here is the obvious fact that the problem method forced students to discover for himself and this forced him into the library, a fact which called for better and better library resources.

In reality the college library, as we would define it today, did not emerge until a sufficient body of information, in active use, was accumulated to require systematic acquisition and organization, and the need for special facilities and staff to perform these services was recognized. Book collections maintained by colleges were simply that until their use became an established, formally recognized activity of the college.

> Lecturing was still popular, but there was less demand for it as students took over more responsibility for working up a field through their own reading. Instead of competing with the printed page as a purveyor of information, the lecturer took on the function of interpretation. The more emphasis that was put on the student's mastication and digestion of his reading, the more popular became the discussion method. Its nineteenth century antecedent, the old fashioned recitation in which the teacher cited and the student recited, went into the pedagogical discard. But the nineteenth century recitation which savored of the Socratic dialectic flourished and took on greatly enlarged significance. (10, p. 275)

The general pattern of higher education which has emerged in America in the present century is remarkably different from that of the colonial college, but diversity of opinion on the philosophy and methods of conducting higher education persists.

While it may be more difficult to see the revolution in college curriculum which has taken place the recent decision to make higher education a right of citizenship and the concurrent burgeoning enrollments are well known. Whereas many colleges a century ago struggled along with 150 students, today there are numerous institutions with over 10,000 enrollment. Rudolph makes the following comparisons:

> In 1870 American institutions of higher learning enrolled somewhat over 50,000 young men and women; a hundred years later the City University of New York alone would be enrolling almost four times that number. In 1870 but 1.7 percent of the young people aged 18-21 were enrolled in colleges and universities; by 1970 half of the age group 18-21 would be at college. In 1960 approximately 3,500,000 young men and women attended institutions of higher learning; by 1970 that figure would be doubled. In 1876 there were 311 colleges and universities; in 1960 there were 2,026. (49, p. 486)

Not only has educational opportunity been broadened (from the 4% of college age attending colleges in 1900 to nearly 50% at present) and been made a part of public policy to provide educational opportunity to the limit of individual ability; but the curriculum has been broadened and the concept of mental discipline has been replaced by the cultivation of intellectual and personal faculties. This cultivation is considered to be a product of intellectual technique rather than specific subject matter. Under such a system education still involves the acquisition of knowledge but more importantly it is a process for the acquisition of knowledge. Moreover, and fundamental to our current library crisis is the conviction that knowledge is considered to be open-ended and evolving rather than fixed.

The open-ended nature of knowledge has caused the number of degree programs offered and the range of occupational fields included in the curriculum to number in the hundreds compared to the single, prescribed curriculum of a century ago. This phenomenon is a reflection of the principle that as the amount of knowledge increases, disciplines tend to fragment. As Daniel Bell indicates:

> A characteristic of science, as of almost all organized human activity, is the increasing segmentation, differentiation, and specialization--sub-division and sub-specialization--of each field

> of knowledge. Natural philosophy, which was an inclusive term in the seventeenth century, sub-divided into the natural sciences of physics, chemistry, botany, zoology, and so forth. Speculative philosophy of the nineteenth century gave rise to sociology, psychology, mathematical logic, symbolic logic, analytical philosophy, and so forth. In any of the fields today, new problems give rise to further specializations: chemistry, which once divided into analytical, organic, inorganic, and physical, is in one accounting sub-divided into carbohydrate, steroid, silicone, nuclear, petroleum, and solid state. (3, pp. 215-216)

The increasing fragmentation of knowledge is preceded by research which not only tends to force specialization upon faculty and researchers but produces a rapid expansion in the growth of knowledge. The rate of growth is evident from the fact that in the United States alone, book production has doubled in theperiod since 1958* and the number of government publications, journals, and reports has increased proportionately. The present volume of information is so great that no one organization can hope to accumulate and organize it for use. The rate of increase is so rapid that no one individual can hope to read even the small portion of new information pertinent to his field. Licklider dramatizes it this way:

> The body of recorded scientific and technical information now has a volume of about 10 trillion alphanumeric characters (i.e., letters, numerals, and punctuation marks) and is increasing along (what for lack of precise data is usually assumed to be) an exponential curve characterized by a doubling time in the range of ten to fifteen years.

---

*American book production was 13,462 for 1958 (American Library and Book Trade Annual, 1960. New York, R.R. Bowker Company, New York, 1960, p. 45.) In 1967 it was 28,762. (Bowker Annual of Library and Trade Information, 1968. New York, R.R.Bowker Company, New York, 1968, p. 88). Both figures comprise new books and new editions.

> To simplify back-of-the-envelope calculations, let us take the figures, 10 trillion characters and (say) twelve years, at face value; let us assume that one thousandth of all science and technology constitutes a field of specialization; and let us consider the plight of a scientist who reads 3,000 characters a minute, which is a rate more appropriate for novels than for journal articles. Suppose that he gathers together the literature of his field of specialization (10 billion characters) and begins now to read it. He reads thirteen hours a day, 365 days a year. At the end of the twelve years, he sets down the last volume with a great sigh of relief--only to discover that in the interim another 10 billion characters were published in his field. He is deterred from undertaking twelve more years of reading by the realization that not only the volume but the rate of publication has doubled.
>
> Sixty years ago, according to our simplifying assumptions, the 3,000-character-per-minute reader needed only twenty-five minutes a day to keep up with everything in his field. Eleven years hence, he will have to read continuously, every hour of every day. (32, p. i)

It is obvious that no student can hope to learn everything he needs to know about his field in the short span of time spent in college. At best, formal instruction can only introduce him to the field, provide a degree of competency necessary to begin practice in the field, and equip him with the skills necessary to continue educating himself. This view of education is reflected in the statement that:

> Few of us think of teaching as the mechanical transmission of knowledge x from person A to person B. In our formal courses, along with information, we consciously transmit a method, a skill, a stance toward problems; and in an age of rapidly changing knowledge, the method, the skill and the stance may well outweigh the information in value to the student. We are sometimes less conscious that in profound ways some of the most

> lasting things we also teach are qualities, abilities, and attitudes exemplified in the way we have taught, in our stance toward the student himself. (26, p. 39)

The impact of changes in higher education on the college library have been profound. The library has become the device for storing and providing access to a growing mountain of information that can be transmitted only partially through formal instruction. Since the growth of knowledge is so great and so rapid, and the disciplines which discover and define it are so dynamic, instruction has to focus on an orientation to concepts and sources and equip the student to study for himself both inside and outside the classroom, and after the end of the course.

The role of the library which has emerged is: 1) the systematic development of collections to reflect the current state of knowledge in the various disciplines, 2) the organization of these resources for prompt access, 3) the maintenance of services to assist faculty and students in the use of resources. Since these resources and services extend beyond the immediate objectives of formal instruction, the college library serves both a supplementary and complementary instructional role for the academic community. As a producer and transmitter of information, academic institutions have generated, at a steadily mounting rate, more knowledge than their supporting libraries can conveniently organize and transmit. There is in sight no conventional means by which libraries can meet the demands imposed upon them.

Chapter II

Technology and the Library

The changing nature of curricular and instructional methods, the increased student population, and the accelerating rate of publication, described in the first chapter, have produced a complexity in library operations which makes it difficult for them to carry out their responsibilities. It is becoming increasingly evident that the money and manpower that could reasonably be made available might still be insufficient, so long as libraries continue to rely on current procedures. New approaches must be found by which some of the emerging technologies, such as computer sciences and microphotography, might be employed to help them in the acquisition, organization, maintenance and distribution of information.

Mechanization of various processes is commonly associated with technology; but in the broader sense, "the science of the application of knowledge to practical purposes,"* is a more accurate definition. Technology is a methodology of resolving problems, and the solutions may be conceptual as well as the application of machines. The approach requires a systematic analysis of the various elements of a problem and an accurate assessment of alternative solutions.

The application of technology to information transfer, has been an area of intensive study for the past two decades. Information technology and information science, associated with this area of investigation, are the blending of disciplines and techniques into the analysis of information problems and the development of new methods of handling information.

The consequence of the vast accumulation of available information is that libraries collect publications in fragmentary and uncoordinated fashion. This, in turn, has led to its inadequate utilization. Individual researchers become increasingly specialized and lose the ability to understand or keep aware of activities in related fields. Considerable time is consumed in locating and using the fragments of information immediately at hand while research effort is duplicated in ignorance of work already completed.

---

*Webster, 2nd Edition.

Much of the accumulated information is under-utilized because scientists and engineers are unaware of its existence. Specialized services such as review journals, abstracts and documentation centers to assimilate and synthesize information have been developed, but these efforts are fragmentary and fail to reach the majority of researchers.

The solution to this problem is not merely in the employment of technological devices. It requires a broad view, and the range of concepts under consideration embraces basic policies, organizational patterns and procedures, in coordination with technological devices.

The solution to information problems is viewed as national and international in scope, encompassing the entire spectrum of production and distribution of information, its systematic appraisal and collection, its organization into files, and its dissemination to users. This approach requires detailed analysis of all elements, not merely the application of machines to existing methods. The systems approach, as this broad view is termed, has attracted the interest of a wide range of specialists: systems analysts, engineers, data processors, photographers, and electronic technicians as well as persons from the basic physical and behavioral sciences.

The information sciences do not promise any easy, quick solution to the problems of libraries. Exaggerated claims which sparked the imagination and marked the early enthusiasm in the field have evolved into a careful assessment of the overall problem and a more precise definition of the basic information required to effect improvement.

The current status of information science, the areas of interest, and achievements are reflected in The Annual Review of Information Science and Technology. (20) The first three volumes of this series evaluate the literature appearing in 1965, 1966 and 1967 respectively together with some of the relevant earlier literature. The contents of Volume I is indicative of the broad range of concerns. Subjects range from the problem of identifying the disciplines which are potential contributors to information science, to the study of the behavior of scientists when confronted with information channels; from the analysis of expressions in natural language, to the problems of library automation; from the history of testing and evaluation of language indexes, to "Selected Hardward Developments."

In describing the structure and content of chapter I Carlos Cuadra, its editor, points out that "it is clear that the overconfident clains of several years ago about high-quality machine translation, automatic abstracting and 'fact' retrieval have given way to painstaking attack on the fundamental problems whose solution is required to support any and all of these applications... (20, p. 11)

The information sciences then, are concerned with the entire spectrum of information from its production to ultimate consumption. Attention is focused on the application of machines to the automatic processing of textual material, to the automatic preparation of indexes, abstracts, and other representations of text, to the formalization of information search techniques, it is focused on the mechanization of access to textual material as well as organizational patterns for implementing changes. The primary motivation for the use of information sciences is to increase the speed of acquiring all of the information which is relevant to a given problem. The approach of information science is to identify functions which can be formalized and are repetitive enough to replicate by machines, to identify elements and processes which require standardization, and to determine appropriate mechanisms for coordination.

Virtually all of the study in information sciences has been directed to the management of large bulks of information. In this respect its application seems more appropriate to the large research library. Why, then, should we apply it to the college library? One obvious reason for attempting to make the attempt to apply information sciences to college libraries is the acute need of those libraries.

In contrast to university libraries and other research centers, the average college library suffers acute poverty of resources. It has been stated that

> a large number of undergraduate libraries in the U.S. lack sufficient scope and depth to provide adequate support of the instructional programs of their institutions. The average number of volumes in the top 60 junior colleges in the nation was 26,620 in 1964 (the latest year for which detailed statistics are available), while the average collection of all 800 was 79,250, the median 54,100 and the lowest 80 collections

> averaged 24,625...
>
> In 1962/63 73% of 4-year college libraries and 91% of 2-year junior college libraries fell below ALA (American Library Association) minimum standards for size of collections.*

In a qualitative sense the meager collections of these libraries are probably more deficient because of indiscriminate methods employed to accumulate the numbers of volumes to meet accreditation standards.

Similarly the staff of the average college library is too small to provide more than a minimal level of service:

> Professional staff members in the most favored 60 junior colleges numbered 3 in 1964. The average of them all was 1.6. The average size of professional staff in the nation's 800 4-year colleges in 1964 was 4.2. The upper 10 percent employed an average of 7.8. The median was 3.0 and the lowest 10 percent of U.S. colleges had, on the average, only one professional librarian.*

The college library, therefore, has an acute need to acquire a sufficient portion of the available publications in order to effectively support the instructional program and it shares with other libraries the need for more efficient indexes and a more efficient access to the documents. These latter two problems are not only common to all libraries but happily enough, are areas which are most compatible to information science.

There are limitations on all libraries which are likely to be overlooked in considering the application of technology to library operations. The options available to the library are severely limited by the format of the information and the problems inherent in its organization, factors over which the library has no control. Many of the proposals for mechanization relate to processes which precede, or could precede, library services. Ideally the manuscript of an author should be converted to machine-readable form so that editing, analysis, processing and other handling could be performed by the use of machines. The publication might appear in the same format or in some other form (such as book, monograph, etc.) which would optimize its use, but the machine-readable version would enhance the utilization of computers and related equipment. Individually no library can afford to compensate for inadequacies of the medium used for dissemination of information; for instance, if documents are not available to it on microfilm, it cannot afford to convert its collection to microfilm. Large research centers such as NASA, AEC or large industrial firms may exercise control over the publication medium it uses and be in a position to take advantage of mechanization. This advantage is not available to most libraries because they are information handlers not publishers.

---

*Paper prepared by Melville J. Ruggles, (Council on Library Resources, Inc., June 8, 1968), p. 1.

*Ibid., p. 2.

The library must select and process information in whatever form it is available. Indexing is the primary device used to integrate information, but, even here individual libraries have little to say about their indexes because, in order to reduce cost, the majority of libraries accept standardized indexes and are in no position to refine or improve the product. Individually, therefore, libraries can introduce only minor refinements over present methods. If basic improvements are to be made, for example in indexing, some approach outside the present framework will be necessary.

The primary objective of most libraries would appear to be increased efficiency in providing reader access to information at the lowest possible unit cost. A first step in achieving this objective is systems analysis, namely, a step-by-step analysis of existing procedures resulting in a comprehensive statement for part of all of the college library's procedures. Such analysis provides the frame of reference for systems design, an objective method of assessing various approaches to accomplish the set of objectives under consideration.

The general steps in library operations which need to be subjected to systems analysis and design are (1) selection, (2) bibliographic search, (3) ordering, (4) receiving, (5) cataloging, (6) processing, (7) file maintenance, and (8) circulation. All appear to be amenable to systems analysis and design.

Only in the process of systems analysis and design is it possible to ascertain the nature and extent of repetitive operations and the applications of technology. (Technology does not imply that a particular procedures may be clearly prescribed.)

Accordingly, the nature and degree of machine operation in the college library setting, if warranted, would be prescribed by systems analysis and design. The application of machines ordinarily requires a high volume of repetitive operations, which in turn could be standardized and result in lower unit costs.

While the college library could gain some efficiency in applying the techniques available in systems analysis and design, it is unlikely that such devices such as computers or microtechnology are warranted in the individual college library. This is because (1) more expertise is required than is normally available on the college library staff; (2) a relatively large expenditure of money is required to exploit this type of technology; (3) the return on investment is likely to be small in view of the small size of most college libraries.

Thus, it appears that technology currently offers relatively little relief for college libraries acting independently. The basic question is whether sufficient volume of work can be developed through college libraries acting in concert and whether unit costs can be significantly reduced. This question will be considered in the following two chapters.

Chapter III

The Core Collection

The operations of a small college library afford only marginal opportunities for the utilization of technology. Although many of its procedures can be analyzed into a subset of highly repetitive operations, the number of times these operations are performed is not sufficiently large to justify the installation of mass production techniques, the volume of its business is too small to attract the research and development interests of industry, and the demands made upon its staff are suficiently diverse to discourage any attempt to make systematic in-house analyses of its procedures.

However, when attention is turned from the isolated college library to the manifold of similar installations in small colleges across the nation, the potential benefits of technology become more appealing. The possibility exists that many of the operations are being duplicated a thousand or more times in as many college libraries. A thousand different college libraries may be separately selecting the same book, ordering and acquiring it, making a search for its bibliographic entry, cataloguing and processing it. The question naturally arises as to whether it is feasible to eliminate much of this duplication of effort. Various attempts have been made to automate, on a local or regional basis, some of the manual and clerical procedures of libraries. The set of operations known as processing has been a notable example. The reports of the results of these efforts is that some economies are effected, but the savings are not dramatic and the services have not been significantly improved. While the evidence is not all in, the tentative conclusion might be drawn that libraries must widen the area of cooperative endeavor if real economies are to be achieved or library services are to be upgraded. An intriguing way to accomplish this is embodied in the concept of a "core collection".

This core can be visualized as a set of library resources which any college library ought to have in order to perform its objectives effectively. In negative terms we might say that if any college library lacked a sufficient number of titles in this core, then its instructional program would suffer. There thus appear to be two considerations in viewing a core: (1) There is a class of library resources essential to the college library; (2) this class of library resources is common, or ought to be common, to many college libraries.

The term "core" will be used to signify two separate but related matters; first, the basic list of resources comprising books published in past years; and secondly, the continuing additions of newly produced books to this basic list.

Although the context of the present discussion is designed primarily for the college library environment, there is a significant ramification in the list of college library resources for the university as well. The identification of the core and the development of procedures to exploit its advantages will be advantageous to the university in providing a more efficient and economical means to meet the library needs of its undergraduates.

The logic for asserting the concept of a core for the college library is not new. For instance, the idea was set forth in 1935 by Louis Shores. (50, pp. 110-114) While Shores' analysis is probably not appropriate to the modern college library it provides a basis for some assumptions which we can now make about the concept of a core:

1. The college library should possess a basic collection of a certain size, based on some standard list;

2. New publications should be constantly added to the basic collection; these additions, too, would reflect a standard;

3. Publications should be deleted from the basic collection as they become outdated;

4. Some form of cooperative processing would be in operation with additions being cataloged and classified centrally;

5. A list of holdings would be issued periodically by the central cooperative agency.

Is there a "core" of literature which is basic for many college libraries? The thesis of this report is that such a core exists and can be identified. Three reasons are offered in support of the idea.

First, the question of the existence of a core collection has been raised with many librarians and researchers in library technology in the course of this study, and there has been nearly unanimous affirmation of conviction of the validity of the core concept.

Second, the implication of an analysis of the methods of acquiring and transmitting information suggests there is a core of literature common to each subject in college curricula. This study, detailed in Appendix I, is abstract and theoretical, and only the summary is reproduced here. The primary objective in instruction is to reconstruct the history and current status of a discipline. The usual method is to select "landmark" figures and contributions to indicate the problems with which the discipline is concerned, the observations which were made, and the interpretations given to these observations. Generally the student is given sufficient exposure through reading, lecture, and discussion to gain a feeling for the intellectual problems dealt with, the nature of the evidence employed, the techniques used in observation, and the reasoning processes applied to the data. The object would seem to be to equip the student to gain competence in collecting and interpreting further evidence as a practitioner in the field or at best to appreciate the tenuousness of factual information and the limitations in methodology in order that he may make appropriate judgments in the future.

If this is a correct interpretation of the instructional process, it may be observed that in synthesizing the literature of a discipline, the instructor extracts portions from the entire body of literature. The amount extracted is that, which in his judgment, is sufficient to explain a concept, illustrate the nature of the evidence and the method employed to establish it. He may emphasize the historical approach or concentrate on current problems, and he may use hypothetical data more meaningful to the students. The portions of the literature used may be highly variable from one instructor to another. He may vary his approach for the sake of novelty or some other reason, but he must not fail to convey the essence and the flavor of the discipline if he is to fulfill his mission. There are a variety of instructional devices employed in the transmission of information in college teaching. Some are books and articles written for that purpose, others are extracted from the subject literature, and others are created by the instructor. There would appear to be a core of "classics" or landmarks that every student must understand to attain proficiency in the "language" of a discipline.

In spite of the wide variation in stated objectives of colleges and the differences in style and methods of instructors, a number of factors tend to impose concensus in the content of the curriculum. The majority of faculty are trained in the major universities of the United States. Each faculty member is trained as specialist in his discipline. The discipline represents a body of information, regarded by concensus of the members of the discipline, as significant. In addition there are problem areas, reflecting its current status, which have not been resolved and on which there are varying opinions. A faculty member will endeavor to transmit to his students as much of the information about the discipline as he is able. His primary loyalty is to his discipline. He recognizes that his standing in the discipline is dependent upon the leaders in the discipline.

A college may take a variety of routes in teaching a discipline, but will be limited, in most cases, by the standards expected by graduate schools. A college values the acceptance of its students for graduate work and gears its procedures accordingly. A further measure of a college's success is the acceptance of its graduates by employers and practitioners. All of these factors tend to impose standards of performance on college instruction. Accrediting agencies, as well, enforce at least a minimum level of quality on programs.

The methods of acquiring and transmitting information in college instruction suggests that there is a core of literature common to each subject. Further investigation into the structure of the literature of a discipline may be fruitful in identifying its characteristics and extent.

So far there has been little experimental evidence offered in support or rebuttal of the core hypothesis. As one experiment, the principal author of this report undertook an analysis of a segment of the literature of the discipline of sociology. The description of this experiment and the results are given in Appendix II. To summarize briefly, the experiment was designed to test the efficacy of identifying the elements of a core collection through an analysis of the citations of authors of college texts and monographs. The results tend to indicate that it is not possible to identify the core elements in this way. Whereas a positive result would have lent additional support to the "core" notion, the null result which was obtained suggests only that it is not possible to identify the core elements with any degree of certainty by the use of citation analysis.

Third, the most persuasive evidence for the validity of the core concept comes from the analysis of the requisition patterns of several groups of university libraries, which seem to indicate that a considerable "overlap" exists in the collections of somewhat similar types of institutions, although individual selection policies were pursued. The extent of the "overlap" or "commonality" in acquisitions is shown in the study of five university libraries (American University, Catholic University, George Washington University, Georgetown University, and Howard University) in the District of Columbia where 200 currently published books formed the basis of investigation. (1)

Another study (22) involves six New England university libraries (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut), which also revealed "a high degree of commonality" in the six collections.

A third survey (17) concerns nine academic libraries in Colorado (Adams State College, Colorado School of Mines, Colorado State College, Colorado State University, Fort Lewis College, Metropolitan State College, Southern Colorado State, University of Colorado, Western State College): here again, the results support the hypothesis that there is a substantial base of common materials.

It should be possible to increase the chances of duplications by considering a group of libraries which are alike in their missions. If all had strong liberal arts programs, then one would expect more duplication among the collections. On the contrary, if one is a technical institute and the others are liberal arts colleges, the chances for duplications are reduced. This idea is supported in the Colorado data in Table I; the Colorado School of Mines (institution #5) acquired only 2.8% of the sample of 1206 titles. This sample was drawn from a very general population, all 30,050 titles being published in the U.S. in 1966. It would be expected, of course, that the Colorado School of Mines acquisitions would be rather narrowly focused in the fields of mining and metallurgy in considerable contrast to the other schools in the Colorado system. Column #4 shows how infrequently the Colorado School of Mines acquisitions duplicated those of any of the other libraries.

Table I

Colorado College and University Libraries and Duplication

Institution Code Number	No. of Volumes Added 1/66-6/30/67	Percentage of Sample Acquired	Range of Probability of Duplication
1	4,014	9.0%	.13 - .38
2	9,631	16.1%	.25 - .45
3	35,583	41.3%	.54 - .77
4	12,858	17.6%	.27 - .48
5	4,733	2.8%	.04 - .12
6	12,204	18.0%	.30 - .52
7	10,012	9.0%	.11 - .27
8	67,466	47.8%	.55 - .81
9	4,466	5.4%	.09 - .17

Another major factor determining the extent of duplication is the way in which the population of titles is defined. In all three cases (New England, Colorado, and Washington, D.C.) this population was primarily, if not exclusively, confined to currently published U.S. books.

A more narrowly-defined population should increase the chances of duplication. For instance, the journal, Choice is specifically designed to assist college libraries in the task of selecting books which are appropriate for academic libraries. Consequently, defining the population as books listed in Choice would be expected to raise the amount of duplication since the probability of a given library acquiring any one of the books listed in Choice is likely to be greater than the probability of its acquiring any one of the approximately 30,000 new titles which appear every year in the U.S.

The Colorado School of Mines showing here is probably typical of the overlap encountered when we juxtapose and compare the acquisitions of university libraries with those of liberal arts colleges, namely, a relatively low degree of duplication or commonality.

Once the mission of the institution is identified, definition of the population is made, and then the chances for duplication are a function of two factors: the number of libraries in the system, and the number of volumes acquired each year by each library.

Clearly, the greater the number of libraries in the system, the greater the chances of having duplications. In the Washington, D.C. consortium of five libraries, the expected number of copies per title was 2.24; in the New England system of six libraries, this number was 3.35; and in the Colorado system of nine libraries, the expected number of copies was 3.82.

All other things being equal, it is also clear that the larger the number of volumes acquired by the library, the larger the number of cases in which it will duplicate a volume acquired by another library. This is because the population, as defined, is constant in size, and, therefore, the larger the number of volumes acquired, the higher the probability will be for any one title in the population. The greater the overlap of institutional missions, the more this will be so. This can be demonstrated by the data available. For the Washington, D.C. consortium, Table II shows the number of volumes processed during the test period and the percentage of the samples of 200 titles which each of the libraries acquired. (1) The ranking is the same with respect to both characteristics.

The results for the Colorado data are not quite so clear cut, but they are nevertheless impressive. Table I shows that the percentage of the sample of 1206 which each of the libraries acquired, was strongly correlated with the volume of acquisition, i.e., the number of volumes processed during the test period. (17) It also shows that the probability that a library would duplicate the acquisition of another library is closely associated with the number of volumes acquired annually by the first library.

Table II

Washington, D.C. University Libraries and Duplication

Institution	Number of Volumes Processed	Percentage of Sample of 200 Acquired
American	26,700	57.5%
Georgetown	23,000	50.0%
Catholic	22,800	47.0%
Howard	20,400	37.0%
George Washington	12,500	32.5%

Although the data for the New England university libraries on the number of volumes acquired during a single year is not at hand we do have the book budget figures for 1964/1965; these indicate a good correlation between the size of the book budget (in lieu of number of volumes added) and the chances of duplicating the acquisition of another library, as seen in Table III. (22)

Table III

New England University Libraries and Duplication

Institution	1964/1965 Book Budget	Percentage Duplication
Connecticut	$579,706	51.4%
Massachusetts	$531,000	67.4%
New Hampshire	$184,000	49.5%
Rhode Island	$140,000	41.0%
Vermont	$125,000	34.3%
Maine	$ 72,777	38.1%

Any attempt to define the core library concept will involve a determination of the numbers of books which will be found in a core and the standard list of books which make up that core.

The feasibility, and the desirablility, of formulating a standard list has been demonstrated. *Books for College Libraries** (BCL), containing approximately 53,400 titles, was prepared for the three campuses of the University of California. *BCL* was prepared with the assistance of numerous subject experts to reflect the most useful resources for a basic college library collection.

Although the preface to *Books for College Libraries* (p. vi) indicates that "the danger in publishing a selection list of this nature is that it may be used as final authority rather than as a guide", it is stated later on in the same preface that, "The project was based on the premise that there is a body of knowledge - - the classics, the important scholarly titles, and the definitive works on all subjects of interest to an undergraduate community - - *which sould be in any college library*." (italics supplied) The *BCL* statement continues, "At the same time it was felt that the library could not grow beyond this figure (i.e., 75,000 volumes) without drawing heavily upon the location, goals, and the curriculum of a particular institution."

The sum of the selections in *Books for College Libraries* might be regarded as a core; actually, *BCL* contains a number of titles which are not currently the best materials - - a matter which could procedurally be rectified by re-issuance of the standard list which could drop the titles which are being superseded.

In this same vein, a core for the junior college library is the *Junior College Library Collection* listing 17,500 titles.**

**Books for College Libraries*: a Selected List of Approximately 53,400 Titles Based on the Initial Selection Made for the University of California's New Campuses Program and Selected with the Assistance of College Teachers, Librarians and other Advisors. Prepared under the direction of Melvin J. Voight and Joseph H. Treys. Chicago, American Library Association, 1967.

***Junior College Library Collection*. edited by Frank Bertelan. Newark, New Jersey, Bro-Dart Foundation, 1968.

It is pertinent to mention two other standard lists in areas distinct from the college library but which nevertheless reflect the concept being developed in this chapter. Law Books Recommended for Libraries* is the standard list of resources published by the Association of American Law Schools. It consists of 40,000 titles and is a "cooperative evaluation of the best in legal literature."(16) Each title is identified with a symbol to indicate the size of the library for which it is recommended; in addition, periodic supplementary lists are planned. The Elementary School Library Collection (Phases 1-2-3)** is a standard list of 6,558 titles (covering audio-visual materials as well); certain titles are marked to indicate when more than one copy is recommended for addition. An advantage inherent in the standard list is that although the primary purpose of the list is a guide for acquisitions purposes, "the format might well be used as a book catalog for the collection and so extend library service in many directions." Supplements are issued regularly and at intervals the standard list is completely re-issued in a new edition; in fact, the third edition (1968) dropped 590 titles (from the second) which were outdated, had been replaced by better publications, or were out-of-print.

Books for College Libraries and the Junior College Library Collection as well as Law Books Recommended for Libraries, and The Elementary School Library Collection demonstrate two things: (1) that the job of defining a basic list of resources can be accomplished through the enlistment of subject specialists; and, (2) that the job has to be done only once for many college libraries, that is, individual college libraries need not replicate this time-consuming task.

A third element is also present. These compilations point to other factors worth consideration, namely, the idea of re-issuance of the standard list to provide for the additions and deletions, designating the level of the titles for different sized libraries, and using the catalog of the standard list as a book catalog for the collection.

---

*Association of American Law Schools. Law Books Recommended for Libraries. South Hackensack, N.J. Fred B. Rothman, 1967. (To be in 6 loose-leaf volumes.)

**The Elementary School Library Collection. Phases 1-2-3. Third Edition. Newark, N.J., The Bro-Dart Foundation, 1967.

A core is not to be considered as a fixed, unchanging list of library materials; it would reflect, however, what a concensus of experts deem best suitable for meeting college teaching objectives, with new publications entering and superseding those which were originally included in it. It is to be expected that the basic concept of a standard list will be challenged by occasional disagreement on the inclusion or non-inclusion of specific books.

Assuming that there is a core along the lines of *Books for College Libraries* and the compilations already cited, what would be the factor of market supply? Would the existence of a core have any effect on keeping books in print? The evidence would seem to suggest a substantial commercial interest in this direction. There are a number of reprint publishers* actively engaged in programs of reprinting *BCL* titles which have been out of print; however, some *BCL* titles will not be reprinted because newer publications have superseded them. This serves to emphasize that the core is a changing entity, with titles entering and leaving.

One book supplier is prepared to supply the monographs listed in *Books for College Libraries* either on a selective basis or en bloc. Although some of the monographs listed in *BCL* are out of print, this firm would secure what is in print, what has been reprinted, explore the second-hand market, and then turn to electrostatic copying for those monographs not otherwise obtainable.

The publications listed in the *Junior College Library Collection* can be purchased en bloc from a book processing firm which would also completely catalog and process the books. The cost

---

*Johnson Reprint Corporation's *Books for College Libraries Standing Order Plan* (5 to 10 titles per month), Kraus Reprint Company's *K-T Book Plan* (10 titles per month, but with 250 titles scheduled for 1969), Greenwood Press, Inc. *Books for College Libraries Standing Order Plan* (10-15 titles per month) may be cited. Greenwood plans on having 1,000 titles available by the end of 1969. Micro-Photo has 442 titles from *BCL* available in its Duopage format (*Bell and Howell Bulletin*, October 1968), and University Microfilms in its *Catalog of Out of Print Books for College Libraries* lists about 3,5000 titles which are also represented in *BCL*.

of the publications (not including the cataloging and processing) would be $210,000 for such a beginning collection.

The fact that a core or standard list attracts substantial commercial reprint support is a significant factor for college libraries. It means that monographs listed in the core would, on the average, be more readily available. This would reduce acquisitions costs in "searching" for books which otherwise would be out of print and difficult to obtain. This, in turn, signifies materials more readily available for instructional purposes.

A standard list requires some method whereby newly published books could be added; this feature, it is noted, was present in The Elementary School Library Collection and in Books Recommended for Law Libraries.

Selection is a responsible and onerous job in any library. It demands continuity of attention on the part of subject specialists if this important operation is to be accomplished effectively.

An idea of the number of books which would be subject to addition to the standard list annually may be roughly approximated by the following reasoning. Of the 28,000 books published in the U.S. yearly about half, or 14,000, would be what we might call "eligible" for some serious or scholarly purpose. Of these 14,000 about 7,000 would relate to professional purposes (law, medicine, religion, agriculture) and would include reprints as well. The other 7,000 would be for higher education, with about 2,000 for graduate purposes, leaving 5,000 for potential addition to the college library standard list of resources. This is shown in graphic form in Figure 1.

There are a number of guides for current selection. The one designed specifically for the college library is Choice.* Starting in 1964 under the aspices of the American Library Association, supported by a grant from the Council of Library Resources, Choice

---

*Choice, a Publication of the Association of College and Research Libraries, A Division of the American Library Association. Published monthly. Chicago, ALA, 1968.

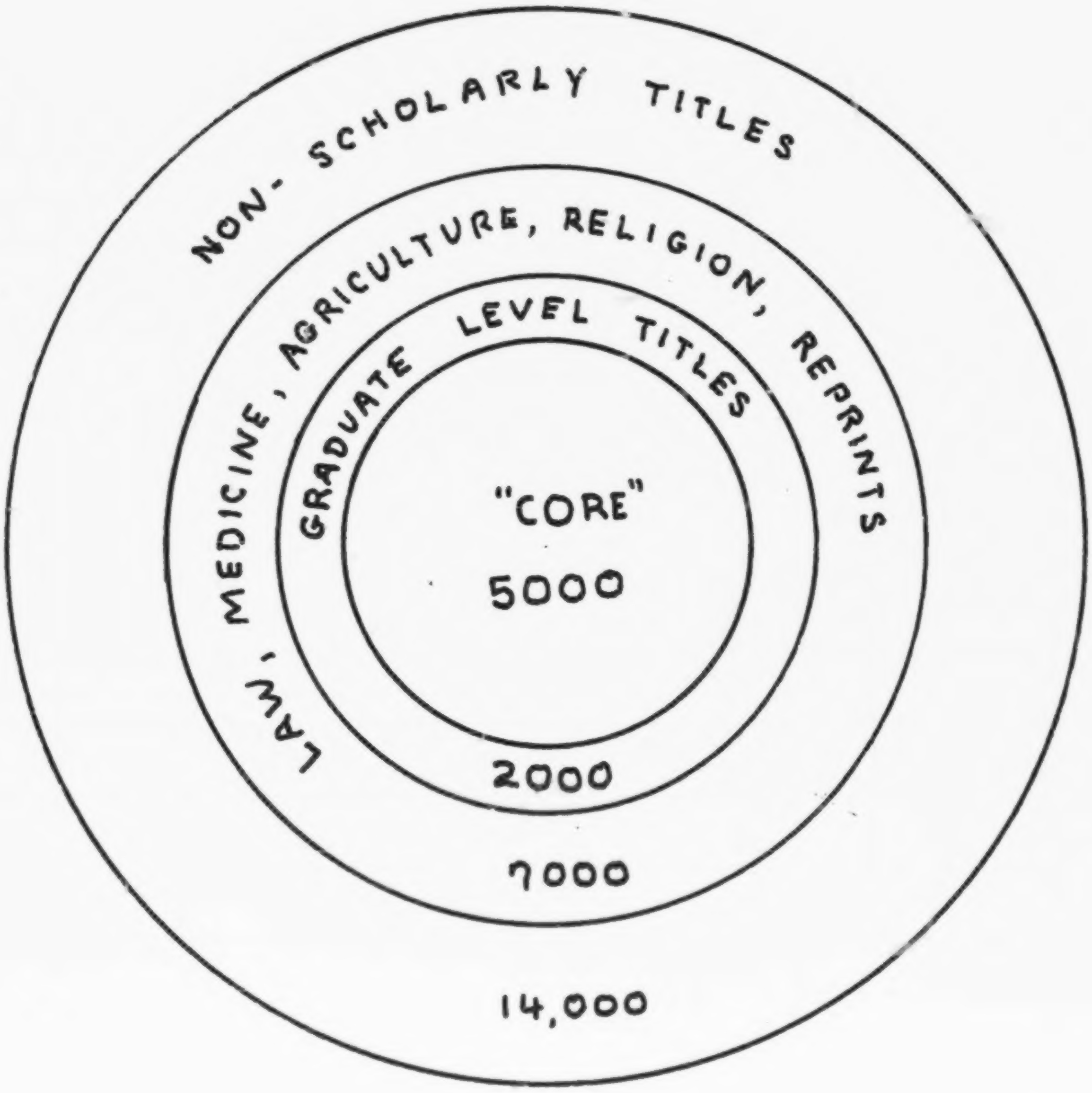

NON-SCHOLARLY TITLES
LAW, MEDICINE, AGRICULTURE, RELIGION, REPRINTS
GRADUATE LEVEL TITLES
"CORE"
5000
2000
7000
14,000

Figure 1

Analysis of 28,000 Annually Produced U.S. Books.

is intended to up-date the selections reflected in Books for College Libraries. Choice attempts to review all current American publications of significance to college libraries with the assistance of 2,000 professors in academic institutions. Selected lists such as the "opening day collection" and "best" books of the year and editorial comments on building collections provide valuable assistance to college librarians and faculty. In addition to Choice there are numerous reviewing media such as the New York Times Book Review, Saturday Review, and scholarly journals. The latter publications generally do not indicate the relevance of the publications which they review for college instruction.

Since it began publication Choice has increased the number of book titles reviewed; this is seen from Table IV which incorporates the U.S. production figures for purposes of comparison.*

Table IV

Number of Books Reviewed Annually in Choice

Volume	Number of Choice Reviews	U. S. Book Production & Year	Percentage of Production in Choice
I	3,388	28,451 (1964)	11.9%
II	4,741	28,595 (1965)	16.5%
III	5,461	30,050 (1966)	18.1%
IV	5,748	38,762 (1967)	19.9%

---

*Source: Issues of Choice for the years noted, Bowker Annual of Library and Book Trade Information, (1965 through 1968 issues) N.Y., R.R. Bowker. Note: Although Choice reviewed 5,748 books, this figure includes some titles which did not receive top-level reviews; thus, a "core" of about 5,000 in order.

However, even with numerous selection guides the individual college library faces a crucial problem in keeping pace with publishing output. Marshalling the subject competence for all disciplines throughout the year at a time close to publication of the books under consideration is an extremely difficult assignment for most college libraries particularly when we consider the level of their staffing. As a result their current additions are often spotty. What is needed is not so much a guide in selection, but regular, definitive, and authoritative selection itself which will be accepted by individual college libraries.

A concensus of subject experts to designate books on a systematic basis is obviously necessary. This could give uniform attention to all disciplines and would avoid the selection task in the college library except for specialized acquisition beyond the core. The time spent in routine selection activities then could be released for more difficult and vital areas of college library service.

Choice, in principle at least, constitutes the vehicle for this job. It is the product of subject experts; however, the listing of books in Choice is lamentably late; the July/August 1968 issue, for example, contains two to three times as many 1967 books as it does 1968 books. More specifically, the delay between the time a library could receive new books and the time these same books are listed in Choice is about 125 days. Desirable as Choice might be in principle as a means for designating the additions to the standard list of resources, this kind of delay makes it impracticable.

A criterion for additions to the standard list of resources or core is that it be available to the college library immediately after publication of the books listed.

The additions to the core would total about 5,000 books annually. While this can be visualized as constituting the list of additions, it would be unrealistic to assume that college libraries overnight could finance such a program. The annual cost of the 5,000 monographs would be about $40,000 at today's prices. It is desirable then to think in terms of levels within the standard list of resources. This amounts to asserting that a Level A book, for example, is indispensable for all college libraries; that a Level B book would be very desirable, and that a Level C book would be desirable, but not as necessary to the college library program as Levels A and B.

Further sophistication is possible within these levels. A given library might purchase all of Level A, all of Level B in the social sciences and literature and all of Level C in literature, for instance.

The concept of the core possessing a certain magnitude implies that the number of selections of new publications would be offset by a corresponding number of deletions of older publications.

To the extent that it is accomplished in college libraries today, "weeding" is usually performed on a sporadic and desultory basis; nor does it keep pace with the intake of new books. The library staff is concerned with selection, as well as the day-to-day operation of the library and little time is left for implementing a policy of withdrawing books, let alone on a systematic basis.

However, the costs of building and storage are such that the "weeding" operation has to be faced squarely from the economic point of view. The steady, cumulative effect of new acquisitions which are not displaced by a like number of deletions implies more shelving, additional floor space, and expensive construction, or over-crowded and inefficiently operated book stacks. It is the crowded stack conditions, and the economic factors, as well as user consideration that point the way to a separation of book collections into different levels of accessibility. (19)

Growth in the university library, with the research mission implicit, may argue for less emphasis on the weeding but the college library does not have the same constraint to retain books as the research library. And it can avoid such costs if a systematic weeding policy could be developed.

If the selection of books for addition to the standard list of resources can be managed once by a group of selectors skilled in various subject fields (just as Choice employs many skilled subject specialists for the act of selection) for many libraries, then it would seem that the same concept could be utilized for "reverse selection", so that monographs are designated for deletion from the standard list of resources once for many libraries.

The college administrator has to confront the consequences of library growth squarely. The withdrawal of books from the library collection is fraught with danger (which is one reason why it is considered "safer" to allow books to repose undisturbed and unused in academic libraries) just as is the analogous act of selection.

The deletion performed by subject specialists would be a persuasive factor for the administrator to consider as part of the core concept. He would have, in effect, the counsel of a group of experts in making decisions periodically on deletions, thereby avoiding or delaying construction costs or additions to the college library buildings.

Procedurally, this deletion can be accomplished by re-issuance of the core which would cumulate the additions up to a certain break-off point in time, by simply dropping the titles earmarked for deletion (just as The Elementary Library Collection meets this issue). Under computer operation, with the standard list on tape, the deletion of titles can be accomplished by dropping them from the new "print-out" of the standard list of resources.

Deleting a book from the college library collection need not mean, however, that the book is irrevocably out of reach. It is simply that the little-used book need not be as accessible as the much-used segment and it can be housed more remotely and less expensively. If the panel of subject experts were to designate books for withdrawal it would be most important that some type of storage facility exist and that retrieval can be accomplished conveniently from that storage facility. This might be managed in several different ways. Rental or construction of low cost cooperative storage facility for several libraries where the withdrawn titles could be kept in compact storage to minimize costs is one possibility. Alternatively, arrangements might be made through a university library or a public library in the area which would function as a regional storage center, and to which the deleted titles could be sent. Not all copies of deleted titles would necessarily have to be stored; some pattern would have to be developed whereby one copy (or several) might be sent to the central storage facility. At the present time the Kansas City Public Library serves as a central depository for higher education (including seven college libraries) in the greater Kansas City area. Under this arrangement, the procedure is to ". . . retain only one copy of each publication deposited, in case of duplicates, and exchange or sell duplicate copies, using the proceeds for the further enrichment of collections to serve the cooperating libraries and their clienteles." (24, p. 40)

It is not expected that a core collection, if it exists, means identical collections in every college. There would naturally be a variety of instructional devices needed for varying levels of students, and this area will undoubtedly vary by faculty preference and style. A further area of literature which is difficult to assess is that described as "source material." These are the materials required for independent investigation and vary considerably in scope depending upon academic requirements. This may be the area of infrequently used material in a collection which needs to be examined more critically for alternative methods of storing and providing access to it.

In summary, the application of technology to the small college library on an optimum basis requires a high volume of materials acquired and processed in multiple copies at the same time. Such conditions would obtain if there were a large number of college libraries whose collections and purchases were substantially the same. This concept, referred to as a "core collection", has not been subjected to a careful, scientific examination sufficiently rigorous to demonstrate its validity. There is a strong concensus among librarians and faculty members that a core collection does exist, and this notion is supported by publishers and book dealers.

On the other hand, in terms of the nature of scientific investigation and the objectives of the instructional process, a much higher degree of common interest might be expected. If an academic discipline can be defined in terms of the objects it investigates and the points of view from which they are observed which, in turn, are derived from common problems and concerns, there is some basis for the definition of its literature. The latter would not be expected to be stable, however, since continued investigation tends to revise and to extend the areas of concern. Evidence of some degree of concensus within an academic discipline is demonstrated by the ability of a widely dispersed group of scholars to contribute evidence and refine interpretations in such a way that independent investigators can reach identical conclusions on the same problem and that the cumulated efforts of the members of a discipline can lead to increasingly refined explanations of their common problems. The explanation that a discipline, such as economics, is what economists say it is seems oversimplified, but syntheses of academic disciplines such as college texts tend to indicate that definitions are reached through concensus.

The impact of standard lists such as BCL and their influence on selection patterns and publisher response in reprinting out-of-print titles lends credence to the concept. The most objective evidence available is demonstrated in the overlap studies in Colorado, New England, and Washington, D.C., cited previously. These studies vary in methodology and are not comparable; furthermore the samples tested are not structured so that subject areas which might reasonably exist in common can be analyzed. The building of college library collections is subject to numerous pressures and constraints such as limited budgets, the uncertainty of whether to emphasize current or retrospective materials, limited tools for evaluation of publications, the lack of expertise on the staff, and uneven and inconsistent attention to selection. Under such conditions it is surprising that collections indicate as much overlap as available evidence seems to suggest.

Chapter IV

Approaches in the Use of the Core Collection

If the premise is accepted that there is a common core of information for college libraries the question arises how could technology be applied and what results could be anticipated? Certainly, continued reliance on the "go it alone" approach in which the individual college library performs all of its selection, ordering, cataloging and processing "in house" is unreasonable and nullifies the advantages inherent in the core concept. Obviously some form of concerted action is required in order to capitalize on the standard list idea. What are the approaches?

One choice might be a Centralized Computer Bibliographic Store. It serves to provide acquisitions, cataloging, and processing data from a central computer store to the member libraries. Under this system the selection remains in the hands of the individual library and little advantage is taken of the core concept. The center pools bibliographic information but does not order, receive, catalog, or process the books; these functions are handled by the individual library which uses the data on hand in the central computer store.

The base of cataloging data is constructed in machine-readable form from: (1) catalog data in processing new acquisitions in the member libraries; (2) from Library of Congress MARC tapes; and (3) from reclassification projects in the member libraries. Teletype access to the central computer from any of the participating libraries initiates a search of holdings to ascertain if the cataloging information for a particular book exists. If such information is at hand the purchase orders are typed automatically at the participating library and when the book is received at the participating library the central computer can provide the necessary data for catalog card sets, labels, book spines, and book pockets. The regional computer center, then, can be used as a tool to provide processing services, but it would have the capability to produce accessions list, book catalogs, and a circulation control. The computer processing center for the New England State University Libraries is a center of this sort. (22)

Another approach might be Processing Center where, on behalf of a number of college libraries, books would be ordered, received, cataloged, and processed by a center and then distributed to the member libraries. Such centers are quite common among public and school libraries but rare among college libraries. The computer can be added to the processing center and would play an important role in facilitating operations and providing certain advantages as well. The processing center can exploit the advantages of the core concept through multiple copy ordering, receiving, cataloging and processing. This would provide an important economy as compared to the "go it alone" technique but it still falls short of exploiting the core idea to the fullest measure possible.

If selection is accomplished automatically then an additional saving is provided. Selection in the individual library can be significantly reduced, with a high percentage of the additions selected from the standard list. Ordering the books constituting the additions can mean the establishment of a continuing arrangement with a book supplier whereby all copies of the designated books are routinely sent to the center. Under this "automatic approach" two advantages would be realized: (1) ordering, cataloging, and processing would be eliminated in the individual library; (2) selection activity for the additions could be largely eliminated and better quality selection provided. (Selection would continue in the individual college library for the non-core titles desired.)

The processing center can be computerized to provide for the preparation of purchase orders, cataloging data, and labels necessary for processing. This involves a file at the center, utilized by the computer, which contains a record of all transactions. The source for this file could be the Library of Congress MARC tapes, with access by LC card number.(2)

It should be noted that "access by LC card number" does not necessitate a conversion of all libraries to the LC system. Such a requirement would be prohibitive and perhaps 70% of college libraries still retain the Dewey Decimal system. Fortunately, the MARC tapes, while their access is by LC card number, can be used to process books by the Dewey Decimal system.

The complete record of additions to the core, along with non-core titles could be annotated by the computer to indicate holdings of the member libraries; in this way the same book catalog could potentially serve a number of college libraries.

The various approaches are presented graphically in Figure 2, and are then described in greater detail.

The subject of centralized and cooperative cataloging and processing is an extensively treated one; more than 900 reports, articles or other published items have appeared, 70% of them since 1950. (31)

A definition of centralized processing (which assumes that individual libraries perform book selection) may be stated in the following terms:

> Centralized processing may be considered to be those steps whereby library materials for several independent libraries, either by contract or informal agreement, are ordered, cataloged, and physically prepared for use by library patrons, these operations being performed in one location with billing, packing, and distribution to these same libraries. (37)

The processing center is an agency ordering, receiving, cataloging, and preparing materials, for two or more libraries. Books and other library materials ordered for cooperating or member libraries follow a logical routine:

1. Delivering the materials to the center by the supplier.

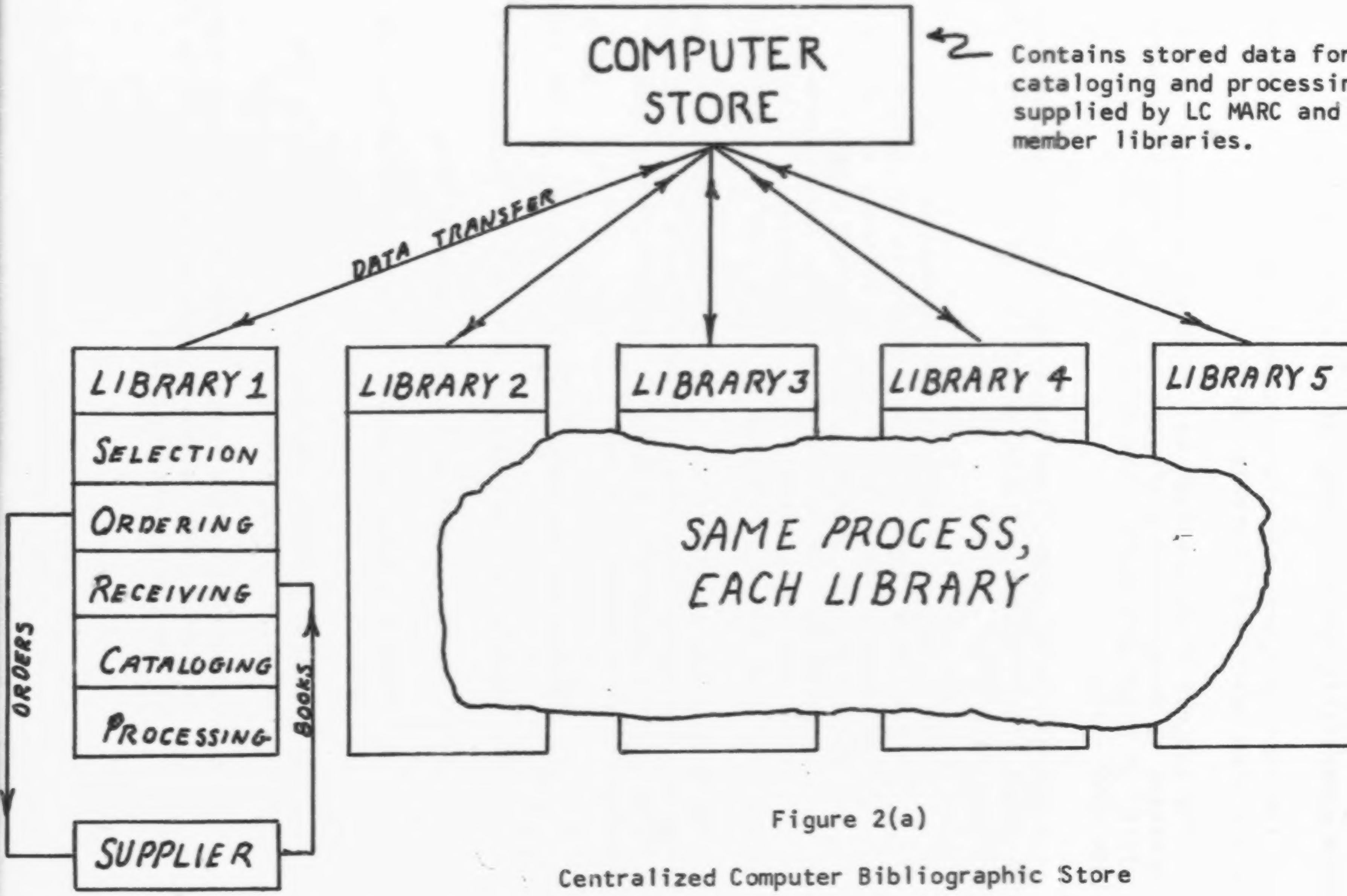
2. Checking and validating invoices for payment.

3. Classifying and cataloging the materials.

4. Preparing catalog cards, pockets, book cards, and spine labels.

5. Pasting pockets, marking ownership, marking the spine, applying plastic covers.

6. Delivering to the member libraries materials ready to shelve, with a set of catalog cards ready to file in the public catalog. (27)


COMPUTER STORE
DATA TRANSFER
LIBRARY 1
SELECTION
ORDERING
RECEIVING
CATALOGING
PROCESSING
ORDERS
BOOKS
SUPPLIER
LIBRARY 2
LIBRARY 3
LIBRARY 4
LIBRARY 5
SAME PROCESS,
EACH LIBRARY

Contains stored data for
cataloging and processin
supplied by LC MARC and
member libraries.

Figure 2(a)

Centralized Computer Bibliographic Store

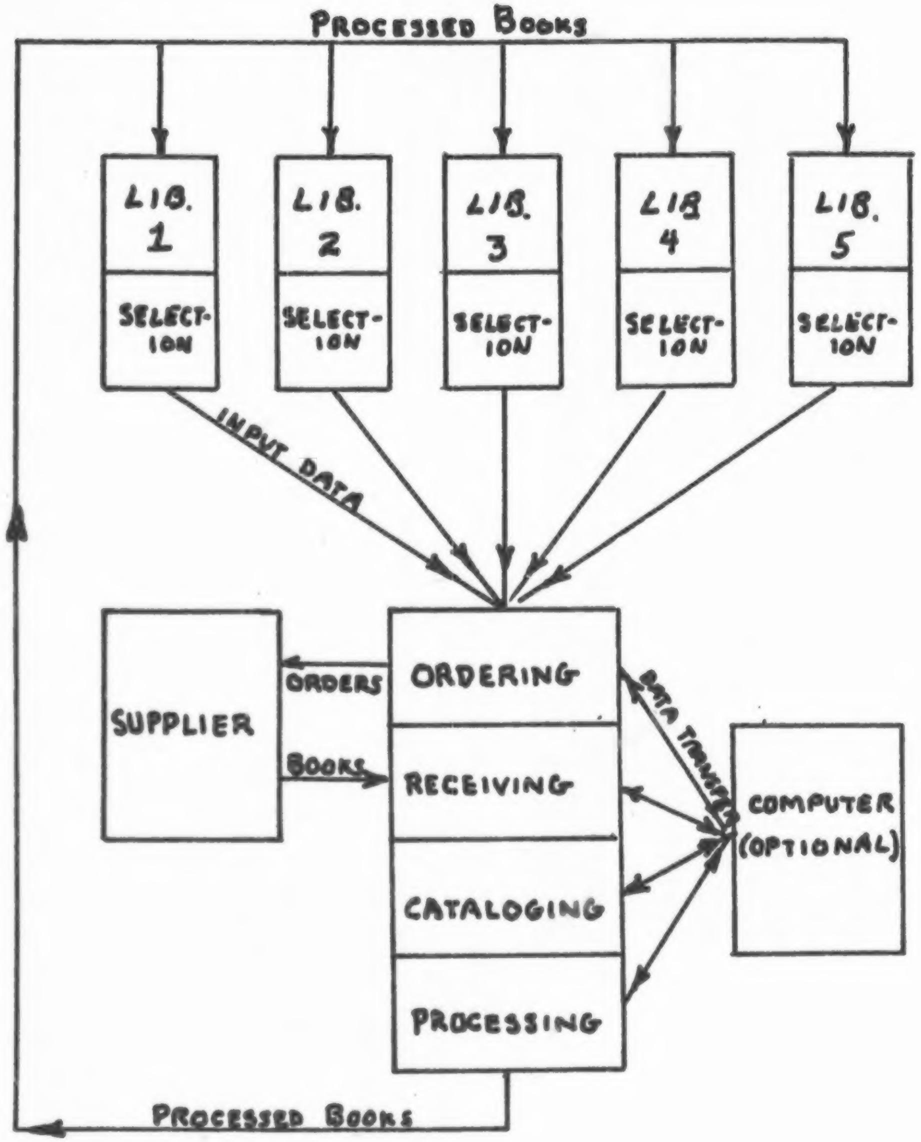

PROCESSED BOOKS
LIB. 1
SELECTION
LIB. 2
SELECTION
LIB. 3
SELECTION
LIB. 4
SELECTION
LIB. 5
SELECTION
INPUT DATA
SUPPLIER
ORDERS
BOOKS
ORDERING
RECEIVING
CATALOGING
PROCESSING
DATA TRANSFER
COMPUTER (OPTIONAL)
PROCESSED BOOKS

Figure 2(b). Processing Center with Individual Selection

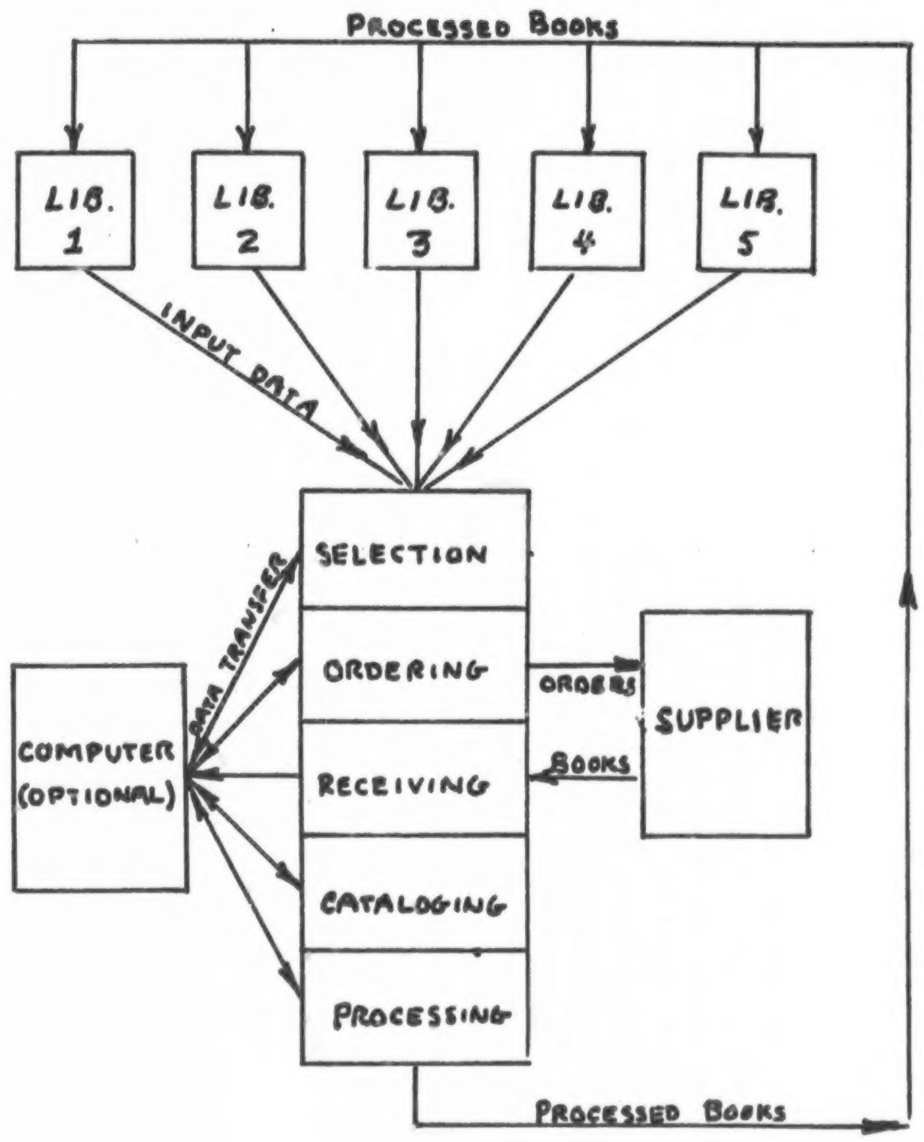

PROCESSED BOOKS
LIB. 1
LIB. 2
LIB. 3
LIB. 4
LIB. 5
INPUT DATA
SELECTION
ORDERING
RECEIVING
CATALOGING
PROCESSING
DATA TRANSFER
COMPUTER (OPTIONAL)
ORDERS
BOOKS
SUPPLIER
PROCESSED BOOKS

Figure 2(c). Processing Center with Automatic Selection based on additions to core

While we lack up-to-date information on the number and characteristics of processing centers in existence in 1968, we do have some data which provide an order of magnitude. Some 2,000 independent public libraries belonged to a regional processing center, according to a 1966 survey made by the American Library Association's Regional Processing Committee.

The median number of member libraries served by a regional processing center is thirty. Annual budgets for operating expenses range from $4,000 to $450,000, with the mean being $142,200 annually. The average cost per volume processed was $1.41 and the average number of volumes processed per year is 56,900 (comprising approximately 9,000 titles). (30)

With respect to school libraries, there were in 1960-61, 467 school systems providing central processing of library materials for elementary schools and 239 for secondary schools; this represented three percent of the nation's school systems for elementary, and two percent for secondary according to the 1960-61 school library statistics from the U.S. Office of Education. (21)

The salient reason for the emergence of the processing center and its spread is made clear from the following:

> Yet another reason for central cataloging for school libraries is the large amount of duplication in collections from school to school. While school librarians do select materials to fit local school needs and a unique student body, they also select to support a common curriculum taught in all the schools of a system. *This duplication of collections increases the possibility of economical central processing schools*. (21) (italics supplied)

Although the statement concerns school libraries the reference to duplication of collections is relevant to junior colleges and colleges as well. The advantages in processing centrally consist in: (6, p. 2)

1. Elimination of needless duplication of expensive cataloging aids.

2. Maintenance of an experienced staff of specialists.

3. Reduction of clerical personnel by using machines for volume production.

4. Saving of professional personnel's time in individual libraries for performance of other tasks.

5. Easy maintenance of a union catalog of system acquisitions.

6. Higher quality of processing.

7. Better discount in quantity procurement.

The anticipated duplication of the order for a given publication is crucial. The Colorado survey describes this factor as follows:

> The expected number of duplicate copies of a title that will be ordered within a system is also an important consideration. Assume that an order is initiated by one library in the system; how many additional orders for that title could a center expect to receive from the other libraries? This is a critical question for any group of libraries contemplating centralized processing, as the relative success or failure of a center will hinge to a large extent on the positive degree of response; i.e. what percentage of title duplication can be expected? (17, p. 28)

Centers may be established to serve groups of public, school, college, or other libraries; consideration is usually given to various factors including:

(a) Size of libraries in the group.

(b) Volume and type of materials to be cataloged.

(c) Cataloging policies acceptable to the majority.

(d) Geographic location of the center.

(e) Transportation facilities.

The administrative organization can reflect State-operated centers; a voluntary non-profit tax-exempt foundation, a voluntary association of libraries; a library system offering a choice of various services to member libraries. To this may be added the commercial service which in its effect provides the benefits of centralization. (27, p. 234)

Consideration in the establishment of a processing center need not rule out the idea of participation on the part of different kinds of libraries. This is brought out in the following comment:

> One decision must relate to the type or types of libraries to be included in one program; another to the standardization of policies and procedures. In formulating plans the presumed influence of variables such as milieu, size, function, budget, and clientele of each library must be weighed against duplicative patterns and standardization which must ignore variables. Among the alternatives for action within a geographic area: (1) to create one processing center to serve all types of libraries; (2) to create a center with divisional sub-structure by type of library; (3) to create centers by types of library; (4) to create within a state one cataloging center with strategically located book depots for ordering and final processing. (52)

The processing center may be one serving college libraries only; however, it is not impracticable to consider a combination center serving both public libraries and college libraries. What might they have in common?

Nelson Associate, Inc., in its study of the centralized processing within the public library system of New York State pointed out that:

> Neither are the dissimilarities between the processing required for the public libraries and pri-

> vate colleges such as to preclude a public library processing center adequately serving the private colleges. A systematic sample of titles from a recent year's issue of Choice when matched against the catalogs of the New York City public library systems showed that 98% of the titles in the sample were cataloged by the system in the same year or in another year. (43)

The reason for existance of a processing center stems basically from the idea that multiple copies of the same book can be more economically handled (ordered, cataloged, processed) simultaneously than by the college libraries acting separately. If selection is based on a standard list of resources then all copies of many books would be handled simultaneously, maximizing the concept of the core and of the processing center. Ordering non-core titles through the center is not as economical as simultaneously ordering several copies of the same title; nevertheless, it would be appropriate to have such non-core title ordering transacted through the center with the bibliographic information clearly spelled out by the individual library. Having the individual library transact part of its acquisitions through the center and another part through non-core titles its own facilities would be expensive and wasteful. The costs of such a fractional operation would be very high indeed, since some staff would have to be maintained in the individual library.

One commerical processor already has a feature of standard list or core ordering through a subject thesaurus. The library simply designates the areas in which it collects in depth, the type of book and level of publication as well as other criteria including the designation of publishers; the budget is then specified and the date for the beginning of the service is set. The commercial processor does not designate books, sight unseen, prior to publication for the individual library but orders well in advance of publication a sufficient stock of any potentially valuable book for the library.

It will take further research to ascertain to what extent the use of a subject thesaurus by a library (or group of libraries) would approximate what would be listed in Choice or in a list of additions to the core selected by a panel of subject experts.

In fact, an intriguing and purposeful investigation would be to ascertain with what degree of precision college libraries and subject experts in collaboration could draft a set of specifications, with the aid of a subject thesaurus, to "define" the books which would be of value for the college library collection. Can a definition be established for additions to the core so that monographs can be nominated either before publication or immediately upon publication? What degree of precision might be expected? The stakes are quite high. Some 125 days between publication and the review in Choice could be eliminated.

Costs of Processing

Costs in the individual library.

The costs of acquisitions, cataloging, and processing vary from library to library. The relative efficiency of the staff engaged in various tasks, the superiority of one set of procedures and work flow over another, the nature and extent of record keeping, the kind and relative difficulty of the cataloging, the number of vacancies, the interaction of all of these factors contribute to variation in cost averages from one library to another. (17)

Comparison of costs from one library to another is qualified by two other elements; the geographic region (introducing variation in labor costs) and the year in which the costs were taken (an inflationary effect).

Although a number of variables in processing routines affect the final product, some idea of the costs to process a volume (exclusive of the cost of selection) are reflected in a number of studies of academic libraries. The University of Denver reaches a figure of $4.33 (in 1960); (56) a figure of $3.76 to $4.33 is computed for college libraries in the Los Angeles area (1962); (40) the California State Colleges a figure of $3.75 (1966) (12, p. 33) and the Colorado colleges average $4.50 (1968). (17)

A penetrating analysis of costs is offered by Brutcher, Gassford, and Rexford (11) with respect to the ordering and cataloging operations; the model they prepared is arranged by level of cataloging difficulty and takes into account printed card availability. In addition, the kind of material (government document, fiction, non-fiction, foreign language book) is considered, as well as the type of library (public, special, college). Their finding for college libraries is that the cost of acquisition and cataloging of a fiction title is $4.40 and for non-fiction, $7.77 (1964). The SUNY survey indicates that "the current costs of processing a book is above $4.50" (1967) in the individual libraries. (2)

These figures provide at least an order of magnitude for the costs of acquisitions, cataloging, and processing in the individual academic library. They show a cost of roughly somewhere between $4.00 and $8.00 per book. None of these figures, however, include the cost of selection.

Costs in the processing center.

Without fully exploiting the cost advantages inherent in the standard list of resource, the available evidence seems to suggest that centralization of the processing operations (without computer) can yield a minimum savings of 25% in cost per volume. The evidence is afforded by the Colorado survey; in the survey the current (manually) individual library cost was $4.50. Under central processing this study estimated the cost would be $3.10. This is shown in Table V.*

Table V

Individual Library Costs Compared to Center Operation

	Acq.	Cat.	Supply	Overhead	Transport	Binding	Total
Ind. L.	$1.412	$1.979	$.292	$.544	$.032	$.367	$4.50
Center	$1.365	$ .981	$.292	$.310	$.057	$.092	$3.10

There are certain processing costs which have to be performed in the individual library. These are estimated at $.59 per volume. Thus the average cost under central processing of one copy per title, is estimated at $3.10 plus the $.59 or a total of $3.69.

---

*Derived from p. 76; Figure 5.7, pp. 125-131; Figure 5.8, p. 132 of Reference 17.

It should be emphasized that these estimates cover only one copy of a given title.

The Colorado study produced a computer model for a book processing center in order to study the behavior of the proposed system under a variety of alternative conditions. One simulation, with the constant being 130,000 copies processed per year, showed that for only one copy, the cost of processing would be \$2.96,* but if two copies were processed at the same time the cost would be \$2.58 per copy, with six copies processed together the cost drops to \$2.33 per copy. (This is a savings of 21.2% per copy when six copies are processed as compared to one copy.) It is the second copy which represents a "payoff", bringing about a reduction in average cost of copies ordered at the same time. Ordering, say, the third copy at a later time means that it incurs the cost as though it were a unique title.

The SUNY libraries estimate a savings for processing of as much as \$2.50 per volume when its center is in full operation, which would be about 50% over the present volume cost. (2) To the cost per volume figure in Table VI (17, p. 199) must be added the cost of processing operations retained in the individual library, calculated at \$.59 in the Colorado survey.

There is a dollar advantage secured by combining the purchasing power of several libraries in establishing a more favorable discount posture. The Colorado survey indicates that a five percent discount advantage would be a conservative guess; this is in addition to the discount already enjoyed by the individual libraries.

The cost to the supplier of sending one copy is not greatly different from that in sending ten copies; therefore, the supplier would have the possibility of "passing on" some of this saving. The vendor would probably take cognizance of (1) total dollar volume; (2) the multiple copy ordering at the same time, and quote on the basis of both these factors. The Colorado survey

---

*The \$3.10 cost per volume compares with the computer simulation figure of \$2.96 per volume; the difference is accounted for by different concepts of allocating supply costs and other factors. (Cf. Colorado study, p. 198, footnote.)

Table VI

Variation in Cost per Volume According to
Number of Copies Processed Simultaneously

No. Copies	Cost per Volume	Savings	Savings, Percent	Percent Cumulative Saving
1	$2.96	$ - -	- -	
2	$2.58	.38	12.8%	12.8%
3	$2.46	.12	4.7%	16.8%
4	$2.39	.07	2.8%	19.2%
5	$2.36	.03	1.3%	20.2%
6	$2.33	.03	1.3%	21.2%

figures the average net cost of a current U.S. book at $7.69 and estimates that group buying will provide an additional average discount of 5% or approximately $.38.

Charges set by a commercial processor for the first and second copies of a book suggest that there is a fifty percent saving in handling the second copy as compared with the first copy, provided it is ordered at the same time.

Savings in processing time.
Another saving, difficult to express in dollar terms, occurs in the total time of processing. It is fair to conclude that substantial savings in time would be possible under center operation. The State University of New York processing center anticipates halving the average time between ordering a book and placing it on the shelves. The time savings arise principally from the rate at which items are shipped from the vendors to the center and from streamlining theprocessing activities by using high efficiency methods, and from changes in the accounting operations.(2) In the Colorado survey the time to process a book in the individual library (the average of six college libraries) is 195 days; the center is expected to improve on that by 25%.

Difficulty of precise cost comparison.

Although we can make a fairly good finding to demonstrate that the processing center is less costly than the individual library approach, the facts and figures necessary to compare the latter with the "coordinated approach", are not clearly focused.

We can, however, write an equation for the total unit cost for a monograph which would cover either (1) the individual library, for which we have some cost figures except selection, (2) the coordinated computer center, exemplified by the New England Board of Higher Education approach as well as the District of Columbia Consortium; and (3) the processing center without computer, for which we have figures from the Colorado survey.

This equation is as follows:

$$TUC = P + T_1 + T_2 + S + O + (L_w \times W) + (Cn - w) + (FC/V) + (A/V) + R_u + Sort_u$$

Where TUC = Total Unit Cost

P = Price per book

Comment: Because of its large volume a processing center may obtain a lower average price per book than would the individual library or the coordinated computer center. This would, of course, tend to lower TUC.

T_1 = Transportation Cost, Dealer to Center

T_2 = Transportation, Center to Library

Comment: What will be the transportation cost of (a) Dealer to Center plus (b) Center to Library, versus Dealer to Library (under individual approach or under the coordinated computer center)?

S = Selection Cost

Comment: Selection will be less costly if the library uses a standard list of resources, saving faculty and library selection time. In addition, the selection according to the standard list will reduce L_w (direct labor, cataloging) since multiple copies of books can be handled simultaneously.

O = Ordering Cost

Comment: Ordering costs would go down as fewer separate orders are placed under center operation (ordering several copies of a given monograph at the same time.

L_w = Direct Cataloging Labor

Comment: Can the cataloging steps--including the acquisition of LC cards--be performed cheaper at the center than at the individual library? Can the computer do the job more cheaply? Can the coordinated computer center do the job more cheaply? Is the computer method cheaper per book cataloged, for one copy of one book? How much do multiple copies reduce cataloging costs (this would be a function of the number of copies ordered, whether occurring by chance or in terms of standard list)? Will greater volume at the center make for greater productivity in manual operations, i.e., in attaching spine labels etc.?

Cn-w = Direct Cataloging costs, non-labor (supplies, computer rental, if any)

W = Wage Rate

Comment: Are the wage costs per hour (including fringe benefits) likely to be different for comparable skills at the center versus at the individual library?

FC = Fixed Costs of Center, except Administration (building space, equipment, etc.)

V = Number of books processed per year

A = Administration costs (Director, section heads, computer programming support in the center)

R_u = Receiving costs (in the using library)

Comment: Receiving inspection must be performed twice in a center is used--once at the center and again at the library. If the center is not used, only one such receiving inspection is necessary . Billing is also doubled--the book dealer bills the center, then bills the library. The number of invoices involved may be fewer because of duplicate books on the dealer's invoice, and fewer invoices between the center and the library than there would be between dealers and the library. Receiving inspection at the center is also facilitated by multiple copies; on the other hand there is also an extra packing and shipping operation with a center, which is not needed when the library orders directly from the dealer.

$Sort_u$ = Sorting costs (before placing on shelves)

It seems likely that the cycle time from starting to select a book to the time that book is ready for shelving can on the average be greatly reduced perhaps from four months to one month (based on the Colorado study's findings). It may obviate the necessity for a good proportion of costly special rush orders.

The extent of consolidation of duplicate book orders from two different libraries is probably dependent upon some sort of blanket authority to the center. Otherwise, the only way such consolidation can take place is for the center to hold a given book order in a backlog for a couple of months, waiting to see if a duplicate is ordered by another library. Such backlogging wipes out the shorter cycle time advantage stated above.

The cost of selecting materials for a college library is difficult to assess accurately because it tends to be a cooperative effort of a number of librarians and faculty members. The amount of time devoted to the process may vary considerably depending on the interest of the individuals concerned. It is obvious that the process is repeated over and over by the hundreds of colleges throughout the country employing essentially the same selection media. If it is assumed that the combined effort of the individuals engaged in selection is equivalent to one man year at a salary of $8000, this effort is equivalent to 1000 titles at current average prices. If the cost of selecting the 1000 titles was distributed over 100 libraries, the annual cost would be $80.00 and $8.00 if 1000 libraries were involved!

Aside from cost of selection, current procedures tend to produce very uneven quality in the building of college library collections. Lack of time or expertise may result in the purchase of material which has little or no utility, and the students and faculty are denied access to the more significant ideas and information which lends quality to the academic program. It has already been pointed out in Chapter II that "the staff of the average college library is too small to provide more than a minimal level of service".* This condition is sure to be reflected in the acquisitions area.

It may be appropriate to summarize this chapter by sketching, in terms of a hypothetical situation, the possibilities for selection and processing open to the college library. We might assume a system by which a panel of experts designate rapidly and periodically the books to be added to the core. Suppose then that the college library is about to accept this selection system through processing center with other college libraries in the area.

Selection: The college library would accept the monographs designated periodically for addition to the core. Thus, a major portion of the selection load will be removed from faculty and library staff; remaining will be the selection of non-core titles to take care of special needs of the curriculum. We do not know the cost of book selection by faculty and staff; however, the selection time and effort would be saved; this saving would be "released" to meet other purposes of the college library.

Ordering costs, labor: All ordering of newly produced books would take place at the center and would be based, to a large extent, on the additions to the core. The ordering of non-core titles would be handled by the processing center also; hence, the college library would be totally relieved of ordering.

The costs on individual library ordering which we have, to apply to our hypothesis, are not too relevant since many of the present costs of ordering are a necessary consequence of "in-house" selection. Thus, the average labor cost in ordering (Colorado survey) is $1.41 if done individually; in the center it would be

---

*Chapter II, p. 20

$1.37, hardly representing the full savings which could be achieved by automatic selection.

Book costs: The books comprising the additions to the core would total up to 5,000 volumes annually; there would be three gradations in the list to suit the book budgets of various size colleges; however in this hypothesis we assume the college library takes all three "layers". Let us assume that the best price the college library could obtain under individual operation is $8.00 per volume (newly produced U.S. imprints). This would cost $40,000. A conservative estimate is that the center marshalling the funds of all participating college libraries, would achieve a further 5% discount. This saving to the college library would be $2,000.

Cataloging, labor costs: The labor cost of cataloging in the college library is $1.98 per book; in the center the cost per book cataloged is estimated at $.98.

Total processing costs: Here we group together the selection, acquisitions, and cataloging costs, plus overhead, binding, supply, transportation (in the library) and we find the cost $4.50 per book plus the indeterminate cost of selection in the college library.

In the center the comparable cost per book is $3.10, but we add $.59 to cover the cost of operations retained in the college library (checking in books from the center, filing cards in the shelf list catalog, etc.). We then obtain $3.69 per book for the center as compared to the $4.50 cost per book plus selection cost in the college library. Based on 5,000 books the cost under individual college operation is $22,500 whereas the center estimate is $18,450 or 18%. All this assumes there is only one copy of each book.

But the center will be requesting multiple copies of books for the various members of the center; with three copies of each title ordered the cost per book drops to $3.05. This means the individual college library cost to process 5,000 books would drop by $7,250. This is a saving of 32%.

Summary on costs: Recapitulating the estimated savings, there would be $2,000 saved by virtue of the volume purchasing power generated by the processing center. Assuming a total of three copies per title processed through the center simultaneously (a conservative assumption given the existence of a standard list of resources) the dollar savings for the college library in processing would be $7,250. The $2,000 could be used to reduce the book budget, or to improve the retrospective collection. Viewed another way, the college library could buy another 250 books with this saving. The $7,250 saved in processing costs could be "released" to other uses in the library. The indeterminate saving in the selection operation could similarly be released.

Chapter V

Bibliographic Access

The index, referred to by Vickery as the intermediate store between the users and the collection (the ultimate store), (54) is an abstract of the contents of the collection which facilitates the location of specific information. The index is the method for providing bibliographic access to the collection, and bibliographic access constitutes one of the basic functions of a library. In any collection, except one which is smaller than any we are considering here, some method of structuring and condensing the contents is required to avoid the necessity of browsing at random. In this chapter current approaches to this process will be described and evaluated and suggested alternatives will be reviewed. Particular attention will be directed to the implications of technology to this college library problem.

Indexing or cataloging, as this process is termed, is the key to efficient access to the materials in a library collection. The records prepared in cataloging constitute permanent records of holdings which support most of the activities and services of a library. The first step in the process, referred to as descriptive cataloging, entails the preparation of an index entry. The entry states "the significant features of an item with the purpose of distinguishing it from other items and describing its scope, contents, and bibliographic relation to other items... The item is described as fully as necessary to achieve these objectives, but with economy of data and expression." (54, p. 20) Elaborate policies have been developed for this procedure by the Library of Congress and the American Library Association for books, monographs, serials, music, films, maps, and other types of graphic materials. The product of this step is the familiar 3 x 5 card in the library card catalog.

A second step in cataloging is subject cataloging, consisting of analyzing the publication to determine what it is about, the intent of the author, and the major topics covered. This information is translated into subject terminology following a carefully delimited list of subject terms, called a subject heading list, prepared by the Library of Congress. Alternative subject heading lists, or thesauri, have been prepared by other organizations, but Library of Congress terms are generally employed in a college library. Although the preparation of subject heading lists lack a firm theoretical foundation, a number of conventions have been adopted to limit the range of terms employed to keep related information together in an index. The techniques which have been developed to control indexing terms are described in more detail by Vickery who states the objectives of the procedure in these terms:

> The purpose of a controlled description language... are: to economize on the number of symbols employed in the description file; to standardize subject description, so that the descriptions of a subject used by indexor and enquirer are more likely to coincide; to maximize the probability of retrieving all documents relevant to an enquiry and none that are irrelevant; and to provide for specific reference and generic survey the the extent needed by the users. (54, p. 61)

Closely related to the subject analysis of documents is subject classification. The purpose of this step is to assign a classification number (or address) for the document in the file. This number uniquely identifies each physical unit in the collection and tends to bring related information together physically in the file. Standard classification schedules developed by the Library of Congress or the Dewey Decimal classification are generally followed. Since books frequently include more than one subject, the dominant subject is selected, and alternative subjects are listed in the card catalog.

The card catalog, the end product of the cataloging process, contains a variety of approaches to documents in the collection. Typically each item indexed is represented by a set of catalog cards consisting of an author card, title card and subject card. These cards are filed alphabetically in the card catalog to form a dictionary catalog affording multi-level access to the items it contains.

A supplementary record created in the cataloging operation is the shelf list, identifying each physical unit of the documents contained in the collection, such as the volumes or parts, and the numbers of copies held in the collection. The shelf list is filed by classification number, reflecting the file sequence of the documents as they are arranged on the shelf.

A similar record, which is actually an extension of the shelf list, is the serials record. This record provides for documents appearing in successive parts, such as periodicals or serials, affording a convenient way to record new additions of each item, and will frequently contain additional information to assist in ordering and binding this type of publication.

The foregoing describes conventional library practices in cataloging items for the collection. The procedures followed have been developed empirically over a long period of time in the management of library collections. Current practices are carefully prescribed and have largely been standardized by the Library of Congress. Since 1902 this library provides, at minimal charge, copies of the catalog cards it prepares, together with the subject heading lists and classification schedules employed in the process. A copy of the Library of Congress Catalog, and more recently the National Union Catalog, is available to libraries in book form. The latter contains cataloging information and a record of holdings for the major research libraries in the United States. Cataloging procedures for most United States libraries closely parallel Library of Congress practices as a result of the availability of cataloging information. Whereas many libraries formerly modified Library of Congress practices to local needs, there is an increasing trend to avoid deviation to reduce duplication of effort.

Cataloging is not quite as simple as it would appear. The local library will need to adapt Library of Congress cards to fit a variant edition, and there are a wide variety of publication forms, such as periodicals, serials, documents, and reports, which the Library of Congress catalogs only partially or not at all. The number and variety of deviant formats has increased sharply in this century and has increased the cataloging load in local libraries or has created voids in the indexing coverage. Serials and periodicals are cataloged as one title by the Library of Congress but the user is interested in the individual articles. In some cases the Library of Congress prepares analytics for monographs appearing in series, and the local library has the option to analyze the parts or avoid analysis. Separate indexes have been prepared commercially for some periodicals and serials, such as the Reader's Guide or International Index, and the Government Printing Office prepares indexes to government documents.

The result has been that most college library collections are structured by form of publication so that each type (i.e., monograph, periodical, document, report, microform, picture, pamphlet, and so forth) is a discrete physical unit with its own indexing source. The card catalog is no longer a unified index to a collection. The user must understand the discrete parts of the collection and the appropriate bibliographic tools in order to gain access to information. These indexes frequently employ different terminology and arrangement making them more difficult to understand.

The controlled vocabulary of indexes appears arbitrary to many users. The terms are inconsistent or lack the depth and precision desired. Cross references are introduced to assist in the translation of the terms used by the requestor into those employed in the index, but they are not standardized between indexes.

A number of special subject bibliographies and abstracts have appeared to compensate for the inadequacies of existing indexes. In some cases these indexes are better adapted to the needs of the user, but the coverage is much broader than the contents of the local library. Specialized library and information centers, serving business and government agencies, frequently perform a variety of services to assist and guide the user of a collection, (such as abstracting, translating, and preparing bibliographies). A more aggressive pattern of assisting users might include summarizing the information on a topic and selective dissemination of new publications. Such services are generally considered beyond the ability of college libraries. The indexes are fragmented by the form of the material, and only the card catalog corresponds directly to the resources at hand. Bibliographic searching is slow and costly for the local library, and is a frequent bottleneck in the processing of new material. Consequently, considerable training and experience are required on the part of the user to gain real facility in using a collection.

Solutions to the problem of bibliographic access are really national in character or at best require cooperative effort on a larger scale. The intellectual processes are too expensive to be duplicated in each library, and there is sufficient common need to command joint action. Two basic approaches are being undertaken, and they are closely interrelated. One is to improve the quality of indexing by more careful analysis of the needs of users and analysis of language itself. The second is to mechanize the process so that elements which are repetitive and can be formalized are conducted automatically. A variety of approaches have been employed with some success, but there is still not enough known about user habits and the structure of language to introduce anticipated refinements. An analysis of various methods and techniques for indexing information is comprehensively treated by Bourne. (8) The linguistic problems are discussed in detail by Borko who states that:

> The goal of automated language processing is to manipulate language by electromechanical means in order to facilitate the communication, translation, storage, and retrieval of information.... At the present, automated language processing must rely on the use of statistical and syntactical techniques. Statistical techniques are based upon the distributional properties of the individual words or specified phrases in the language. In its simplest form, statistical processing of language consists of counting the words in the text. The principle on which this procedure is based is that the meaning of the document is carried in the words in the text -- that is, the vocabulary. Articles on different subjects use different vocabularies; therefore, if one knows the vocabulary, he should be able to determine the subject content. (7, pp. 5-6)

Borko sees automating the syntactical analysis of natural language as a necessary first step, but additional semantic problems require solution before the quality of language processing can be improved. Simmons' assessment of the current state-of-the-art is that:

> The accumulation of research findings to the present has been neither notably barren nor noted for significant breakthroughs. A significant understanding of dictionary structures, grammars, and parsing algorithms and a beginning of appreciation for some aspects of linguistic-logical inference possible in language have been attained. Several rather large grammars (for

> rather small subsets of nautral languages) have been developed. Automated systems using such grammars for syntactic analysis exist. Reasonable (but still inadequate) paradigms for the process of translation from one language to another have been developed and tested on at least a few sentences or paragraphs of natural language materials. Such aspects or morphology as hyphenation and the stripping of affixes have become well-developed computer technologies.
>
> The limited information and technology that have accumulated have proved useful in harnessing computer power to such basically human tasks as content analysis, text editing, machine-aided translation, and data gathering for linguistic and stylistic analyses. (20, p. 138)

The most immediate hope for libraries in the improvement of bibliographic access is in the use of computers for automation of various processes. An assessment of how this approach is likely to proceed over the next twenty years is given by the Educational Facilities Laboratories:

> Its first general impact will be in the area of housekeeping chores -- order records and reports, fiscal control, circulation systems, etc. Application of the computer to bookkeeping operations, such as buying and receiving, is the easiest to accomplish. Moreover, it is the area in which considerable experience in many libraries to date indicates clear cost advantages in large operations.
>
> The second field of general application--and impact--will be the computerization of the library card catalog. Some aspects of this are now technologically feasible. The promise is that this will extend greatly the usefulness of information contained in the present card catalog. Its advantages lie in its accessibility to users beyond the library, in permitting the interchange of catalog information between libraries, and in mobility within the library itself for checking holdings, changing location records, and the like. While the conversion of any library operation to automation must be under-

> taken with the greatest care and planning because of the very large costs involved, this is especially so with regard to computerization of the card catalog.
>
> Totally aside from costs, there are technological and intellectual problems of the greatest magnitude to be overcome before computerized catalogs will be generally usable. Direct access files of larger storage capacity than presently available in computer systems will be required to store the catalogs of great research libraries. The capability of simultaneous consultation of the catalog by very large numbers of users must be expanded (the present limit is about 30). Problems of what terms and how many to use in describing catalog information must be resolved. Programs to retrieve only the materials specifically required by a user must be developed. But, despite these problems, it is expected that within 10 to 20 years, the use of computerized catalogs will be widespread. (25, pp. 10-11)

While large libraries may be able to computerize its catalog it is apparent that no college library by itself is in a position to exploit computer technology unless considerable money and expertise is available to support the project. The picture has been altered to some extent since the writing of the Educational Facilities Laboratories report, however, and needs to be explored. A standard format for the preparation of machine-readable cataloging information has been developed by the Library of Congress in its project MARC and has been adopted as a standard by the American Library Association. It is anticipated that magnetic tapes for current American imprints will be available for purchase by libraries in the Spring of 1969. A second factor is the emergence of processing firms to prepare library materials using computer technology. An additional factor is the attempt to develop standard programs for the use of the computer by groups of libraries. Experimental work along this line is being conducted at the Systems Development Corporation and other companies, and systems are being developed in New England, New York, and Ohio in the academic library field.

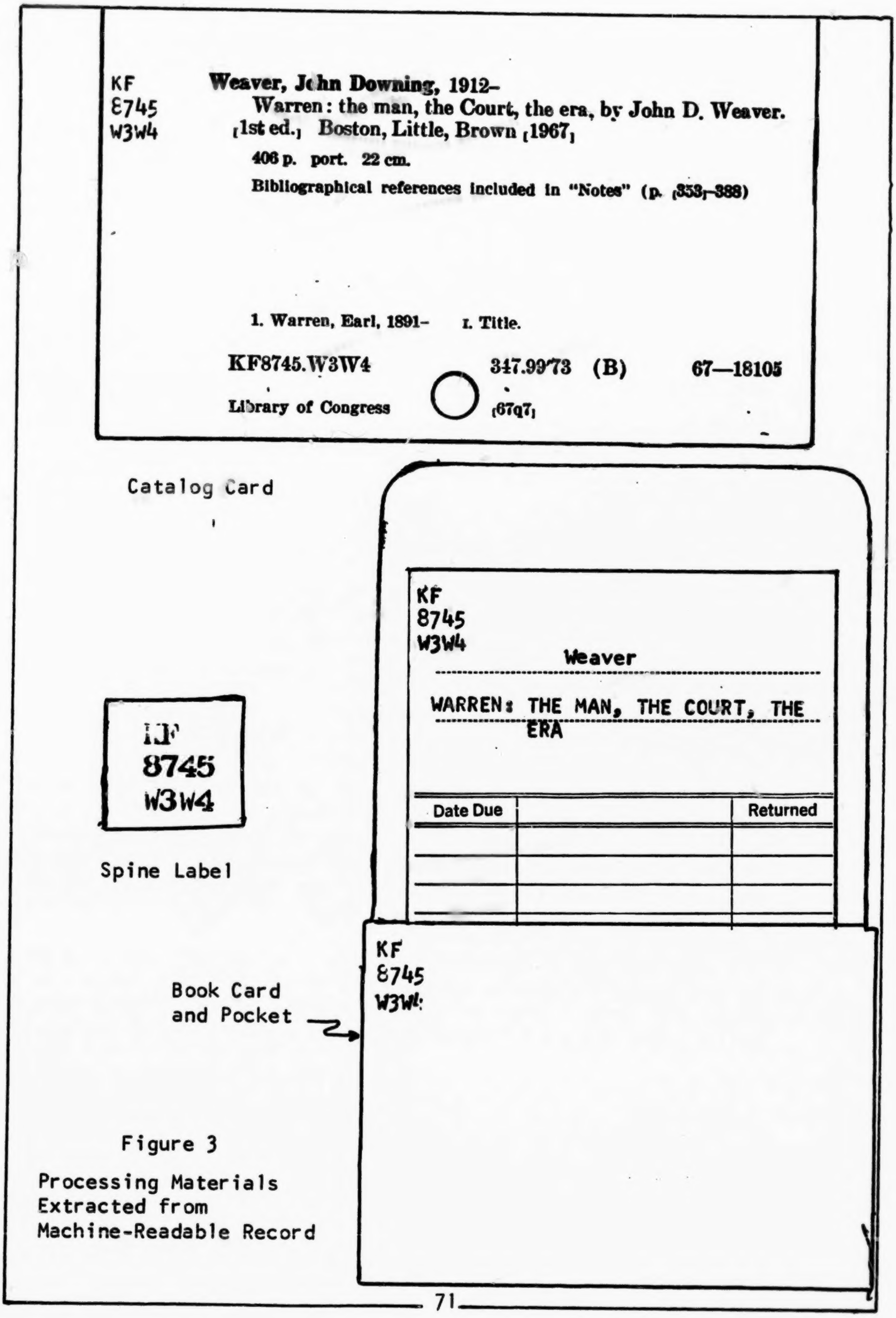
The impact of the computer on the library would be much broader than just bibliographic problems because of the close relationship cataloging information has to other libraty operations. In examining the format of machine-readable cataloging information, it is evident that many of the acquisition, processing, and circulation routines can be mechanized. A simple illustration is to examine a conventional catalog card, Figure 3.

By attaching flags or codes to the various elements as is done in the MARC program, any element or combination of elements can be extracted, rearranged, and reporduced. Hence spine labels, circulation charge cards or book pockets can be prepared by the computer. Additional information can be added to elements extracted from the MARC tapes so that a new record such as a book order can be prepared.

The use of a simple identification number such as the Library of Congress card number can be used to represent the entire Library of Congress bibliographic entry so that the computer could be asked to perform bibliographic searches for acquisitions purposes or for interlibrary loan requests. The same reduction principle can be used to maintain circulation records where the bibliographic entry code and a student identification number provide an adequate record of the loan. In case a book is overdue the description of the book and the borrower's name and address can be extracted and combined from the bibliographic and student record tapes respectively.

A library can inaugurate an automation program at a number of points. Some libraries have automated book ordering, circulation records, or serials control separately. Cataloging information has generally been avoided in view of the MARC project at the Library of Congress. The availability of the MARC information will make it feasible to integrate these operations through the use of the bibliographic code number described above. Ability to integrate processes is important in eliminating repetition in typing. This becomes significant when a library collection is large, and there are a large number of records to be converted. The employment of some code device to tie the records of different processes together will enable a library to build an integrated system as machine-readable cataloging information becomes available.

Library automation is described in more detail in the literature. A convenient starting point would be to examine the



KF
8745
W3W4

Weaver, John Downing, 1912-

Warren: the man, the Court, the era, by John D. Weaver. [1st ed.] Boston, Little, Brown [1967]

406 p. port. 22 cm.

Bibliographical references included in "Notes" (p. [353]-388)

1. Warren, Earl, 1891- I. Title.

KF8745.W3W4 347.9973 (B) 67—18105

Library of Congress [67q7]

Catalog Card

KF
8745
W3W4

Spine Label

KF
8745
W3W4

Weaver

WARREN: THE MAN, THE COURT, THE ERA

Date Due		Returned

KF
8745
W3W4

Book Card and Pocket

Figure 3

Processing Materials Extracted from Machine-Readable Record section on library automation in the Annual Review of Information Science and Technology (20) or the bibliography prepared by the American Library Association. (39, pp. 674-694)

The feasibility of attempting automation in a small library depends on the availability of computers and accessories. Computer availability can also be a detriment which can be misleading, for the existence of a machine on the campus does not mean that the major hurdle in library automation has been crossed. Computer programs are expensive to prepare and are rarely transferable from one library to another. Often, it is easier to start over again than to modify an existing program. This situation exists because almost no two computer installations are identical. Each has its own unique characteristics. Even if the make and model are the same, storage capacity and type, input and output devices and other characteristics may vary. Table VII crudely illustrates the variety of components which are available (listed in ascending order of sophistication).

Table VII

Variety of Computer Components

Input device	storage	computer	output
console typewriter	tab cards	processor	console typewriter
tab cards	paper tape	(vary in	printer
paper tape	magnetic tape	capacity)	Cathode Ray Tube
magnetic tape	disk		
optical character reader	computer memory		

The fact that library automation can be achieved in part with relatively simple components makes it tempting to try. It may be a better investment, however, to take a longer view and have planned an integrated system rather than attempting the project piecemeal. Since this is impossible for the college library it will be necessary to explore various avenues of cooperative effort.

Most of the discussion of library automation and problems of bibliographic access relate more to issues in information systems and large libraries than to typical college libraries, whose collections are more limited. These collections are increasing in size because publication rates are rising, and the scope of educational programs is growing. There is a tendency for requirements to increase in rigor as secondary school education improves in quality. Further, educational programs today do require more extensive library use as instructional methods place greater responsibility on the student. College libraries contain the same kinds of resources and employ the same procedures as the larger research library. The problems may not be as acute, but they are generally of the same character.

There are other aspects of the college library situation that are unique. In instruction, the faculty member frequently knows exactly what he wants the student to read and the sequence in which it is to be read, but the structure of a library collection is a clumsy package for his purposes. The faculty member culls through a wide range of information, making frequent compromises because the collection is not like the one at the university he is accustomed to use, and he finally prepares a bibliography, consisting of a wide range of sources in sequential order for the student to read. A number of sources are omitted because it is easier to summarize or refer to them in class. He has, in fact, designed a custom textbook. The library is placed in the awkward position of attempting to reproduce the professor's anthology on call or any portion of it, thus placing a heavy demand on the library, and representing a substantial part of its work load, for what is basically a clerical, merchandizing operation. The photocopying machine, copyright restrictions notwithstanding, is the only avenue of relief short of buying an excessive number of duplicate copies and which in addition, are often out-of-print. There is, though, an alternative approach to this problem which will be discussed in the next chapter of this report.

The function which the college library is best equipped to serve, adn what in the eyes of many instructors is a more significant pedagogical device, is responding to individual search requests. The problem approach requires the student to imitate the approach of a scholar, to define his problem, examine the evidence, and make an independent interpretation of the findings. It requires the same technique that he will employ in later professional life and gives experience in using bibliographic tools required for continuing education after leaving college. To the extent that this method of instruction is employed, the library may need access to greater resources and services. The student will be placed in the same frustrating position that any library user is in to locate information, but these limitations will eventually be improved. In the meantime the student, through the habit of using informational materials in problem solving, will be able to use the information systems of tomorrow more advantageously.

In general the limitations on bibliographic access to library materials are national in scope. A great deal of research and development activity is underway to alleviate these problems. The immediate approach is through library automation, but this is not an easy avenue for the small college.

Future developments in improving bibliographic access to materials may alleviate some of the internal problems of the local college library, but the more significant developments would appear to come from more effective coordination of responsibility for assembling, organizing, and disseminating research materials required to supplement the college library. As Clapp stated:

> However, an examination of the situation technological developments for sharing and disseminating resources leads to the conclusion that the most useful development of all for the purpose of sharing might be the assignment of specific responsibility for certain conspicuous subjects to particular libraries, each of which would be obliged to acquire comprehensively in the subject, organize and publish its bibliography, and render a nation-wide (or perhaps even international) backstopping lending or photocopy service in the subject. (14, p. 36)

The technology and methods by which this network concept could be achieved will be discussed in the following chapter.

Chapter VI

Physical Access

The printing press and the variety of other devices available for the recording and dissemination of graphic materials has revolutionized our ability to communicate. It has given us such facility to exchange information, thought, and emotion that we have been able to effect near universal literacy, attain a higher level of education, transform the conduct of business and government, and add to the enjoyment and enrichment of life. We have become so dependent upon the exchange of information that accessibility is essential, but the faster, cheaper methods of production and the number of persons with facility for communication have overwhelmed us.

The book, our basic graphic form, has served us well for centuries, since it can be produced relatively inexpensively in quantity in a wide variety of sizes, formats, colors, type size, and quality of paper. A book with its textual and illustrative material can stand as a work of art. Further, its portability and relative durability, have made it the universal medium for information transfer for centuries.

However, the graphic medium is losing some of its appeal for storage and transportation of information in our society. It imposes on libraries a bulk that increases cost and limist accessibility. Libraries might be accommodated to these factors if it were not for the exponential rate of increase in the production of information. Fremont Rider dramatized this situation when he projected the growth rate of the Yale University Library and indicated that by the year 2040 it would contain 200 million volumes and the card catalog along would require eight acres of floor space. (48) Dix calculates that even a collection of two million volumes requires three and one half acres of floor space; and by the time space for readers, staff, non-book materials and other accessories are added, the building for a library of that size would require seven acres of floor space. (23, p. 5) If the predictions of Rider, Ridenour, and others are correct these space requirements will on the average double every sixteen years. It is clear that there are limits to the time, patience, and physical endurance required of the user of a library. Yet these limits are tried by the amount of material he must consult, the wide dispersal of books in the stacks, and the physical separation of indexes from the book stacks.

Number of books in a library alone is not the whole problem but increased use per student compounds it. If we assume that the average student consulted 80 books a year as indicated by Cornell University, (51, p. 72) it is not surprising that a substantial portion of the library's budget is spent on stack maintenance alone. For that would mean that a moderate sized university of ten thousand students would have over twenty miles of books each year to keep circulation records on, replace on the shelf, and keep mended and rebound.

The fact that library materials are serving increasing numbers of students and that we tend to exploit the portability of books is often a source of frustration to other users. The stacks are not kept intact, books require mending, and the complexity of handling bulky materials imposes delays on returning items to the stacks. The degree of accessibility which can be maintained for library resources is considerably below the desirable level.

Consider the limitations on maintaining access to materials in a library together with the considerable and increasing cost of acquiring and processing these materials, it is obvious that other alternatives must be explored. It is evident that, considering the cost of maintaining collections in their present form, we are reaching the point of diminishing returns on the level of service we can maintain.

Attempts to alleviate problems of bulk by sharing the cost with other libraries have not proven very successful. Institutional rivalry and ambition erodes acquisition agreements originally proposed to divide the responsibility for collecting by subject fields. Lack of adequate bibliographic information, complexity, cost, and uncertainty of inter-library loans have tended to limit this form of cooperation to a last resort approach, although the availability of union catalogs for a number of large libraries and the use of teletype, photographic reproduction, and

and delivery service have alleviated some of the problems. Cooperative storage programs have not notably slowed the growth pattern of participating libraries. The insistent demands of productive scholars and the desire to broaden exposure to information for students are not easily met through these methods.

From the point of view of the library, the desirable storage medium for documents is one that is inexpensive to acquire and process, durable, and easily transportable. The library user, on the other hand, requires convenience and comfort in use, and assurance of access in a reasonable time. In some respects these are conflicting requirements, and the problem of physical access is that of finding the media for storage of information which are the most effective compromises.

In recent years, serious consideration has been given to two alternatives for the storage of indexes and documents: the computer and microform. Both of these approaches have several common characteristics: compaction of information, rapid access time through machine handling, and facility for rapid transmission over long distances, in a variety of forms, and at relatively low unit costs. For several reasons, however, they are more interesting for the potential that they appear to have than their immediate availability. Considerably more research and development work is required to make them general utility media.

Computer Technology and Libraries

Although initially designed for numerical computation, the digital computer has demonstrated its ability to handle textual material by the simple expedient of converting alphabetic and diacritical characters into digital codes.

It is essential to realize that the computer cannot recognize and manipulate characters as such, whether alphabetic or numeric. It responds only to mechanical, electrical or magnetic fluctuations. Thus, it can sense the presence and position of holes punched in a tab card, the presence or absence of light signals, or the magnetic variations on a reel of magnetic tape. Auxiliary devices must first be used to convert numbers and letters to hole punches, light signals or magnetized sections of a tape before the computer can be put to use. The process of doing this is called conversion to machine-readable form.

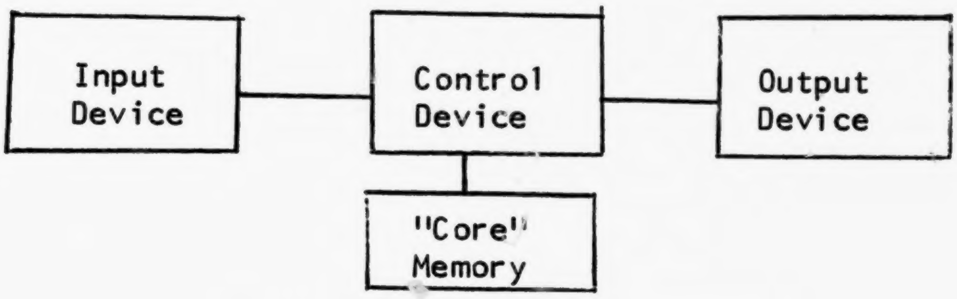
There are a number of ways to accomplish this. By far, the most common method is to utilize a keypunch to convert the information into punched cards or punched paper tape. Keypunching is very similar to typing except that additional codes and symbols are added to direct the computer; hence the amount of typing may be more than simply copying the original document, depending on the process to be performed. It is alos possible to convert information directly to magnetic fluctuations on a reel of magnetic tape by a similar typing process. A more recent development, the optical character reader, converts a document into machine-readable form by scanning the document photoelectrically, character by character. There are severe limitations on its use at present, however, since most optical character readers can respond to only a limited range of type fonts and sizes.

Finally, it is possible to make the conversion directly into a computer's magnetic memory by typing at a console typewriter tied directly to the computer. This eliminates the necessity for the intermediary cards, paper tapes, and reels. However, it is the most expensive way to make the conversion, since it ties up the high speed computer while the essentially slow-speed typing operation is being performed. If the computer is used in a time-sharing mode, of which more is said later, this disadvantage can be overcome. However, it is generally desirable to have the document stored in its machine-readable form, and the intermediary cards and tapes serve this function quite well. Even where the information is converted directly into the computer's memory, it is usually desirable to direct the computer to make a copy of the information onto magnetic or paper tape, on magnetic disks, or on punched cards for storage purposes.

There are certain advantages and disadvantages associated with each method of storage. Magnetic or paper tape must be read sequentially; that is to say, a record near the end of a tape can be reached by the computer only after causing the tape to be unwound until it reaches that section, a process that is slow compared to the intrinsic speed of the computer. Paper tapes and cards are fragile, as is attested by the frequent injunction not to "fold, spindle, or mutilate." Furthermore, cards are bulky, and can hold only 80 characters each.

Storage on magnetic disks offers the greatest accessibility to various parts of a record, since any part can be reached directly without the necessity of scanning unwanted records. Access time to any given record on a disk or drum is measured in tenths or hundredths of a second. But whereas a disk can store about a half-million "words", a 2400-foot reel of tape can store about a million. Furthermore, storage costs per word varies significantly among the storage media, .015¢ to 3¢ per word, hence a compromise between cost and access time may be required. Significant developments are being made in improving storage media and the costs are going down.

Once converted, alphabetic characters can be manipulated in the computer with much the same facility that numerical data is handled. A simplified diagram of a computer is shown in Figure 4.



Block Diagram of a Computer
Figure 4

The input device may be a card or paper tape reader, a tape deck or, as mentioned above, a console typewriter (see Table VII, p. 72). Not only must the information to be operated upon be fed to the computer through this input device, the instructions to the computer must be entered here also. A series of instructions, called a "program" must be written, converted into machine-readable form, and transmitted to the computer before the information can be processed. This is a time-consuming operation and can be as expensive as the data-conversion cost or the direct computer cost. Program writing requires the talents of specialists-programmers--who must make a careful analysis of the operations to be performed and prepare explicit directions. Programs are written in special languages, and these vary with the computer installation, so that programs are not directly transferable from one organization to another. Even after the program is written, a period of extensive testing and modification is necessary to make the program perform the desired operation successfully. The planning, writing, and testing of programs must be considered as an additional expense until the design is perfected to the point of replacing the existing procedure.

The computer proper, consisting of the "control" device and its associated "core" memory, receives the instructions to be performed, stores these in its memory, and proceeds to follow the directions, reading in all or part of the information, manipulating it in the way it has been directed, and disgorging its results to the "output" device. The bewildering variety of computers differ essentially only in the size of their core memories, and in the speed with which they can carry out the program steps. With larger memories, they can store more sophisticated and involved programs, and a larger portion of the information to be processed. The faster they operate, the quicker they will complete a given program and get on to the next. The larger their memories, and the faster they operate, however, the more they cost. And it is not always appreciated, by laymen particularly, that the limiting speed of the computer is determined, not by the intrinsic speed of its control mechanism, but by the mechanical devices at the input and output by which it communicates with the user. Where lengthy records must be fed to the computer, and lengthy printouts retrieved, the prime requirement is on the speed of the input-output devices--high speed readers and printers. A most significant recent development in this area is an output device in which the computer's results are "printed" directly on microfilm at speeds approaching the intrinsic computer speed.

The output device can be a card or paper tape punch, or a magnetic tape recorder (Table VII, p. 72). In such cases, it is necessary to convert the digital codes represented by punched holes or magnetized sections of tape back into printed characters. The most common output devices do this directly. Besides the microfilm device alluded to above, this may be a console typewriter, a high-speed printer, or a cathode ray tube.

The console typewriter is the slowest of theoutput devices, having approximately the same output speed as a fast typist (i.e., about 100 words a minute). It has a character set equivalent to a normal typewriter, and the quality of its output is similar. Its principal advantage is that it can be connected ("on-line") to the computer by cable, telephone or leased lines from remote locations. They are effectively used with "time-sharing" computers, where a number of terminals can be connected to a single computer simultaneously.

The high-speed printer delivers a whole line of up to 150 characters at a time at a rate of several hundred to well over a thousand lines per minute. The quality of the printing is less than that obtained from a typewriter. The printout is generally in upper-case characters only, although printers with upper-and lower-case capability can be acquired at additional expense, and at a sacrifice in the printing speed. It also suffers from the disadvantage that it must be installed adjacent to the computer.

The cathode ray tube terminal consists of a television-like tube in sizes up to 23 inches, and a typewriter keyboard. Like the console typewriter it operates "on-line" with a computer, and has multi-terminal capacity to remote stations. It can display a full screen of print at a time. In more sophisticated models, a light pen can be used for editing the displayed text. Special cable links are required for long distance, and the image quality is not as good as the printed line.

The last few years have seen the emergence of "time-sharing" computer systems. Basically, the notion is straight-forward: since computers react so much faster than people do, the computer can be working on one problem while several others are being prepared for it on the relatively slow input devices, or while previous results are being delivered on output devices, which are also slow. Technically, the problem is involved, but the essential techniques are solved and in operation. Elaborate switching devices place the computer momentarily in contact with first one and then another of several terminals connected to it. Up to 30 terminals can be handled by large computer systems. The terminals can be the cathode-ray consoles already referred to, or even smaller computers. More commonly, they consist of the typewriter consoles with the addition of an internal "memory," like that of the computer's core, except much smaller. An an operator prepares a program of instruction or enters data to be processed on the typewriter, the typewritten characters are converted into their machine-readable codes and stored temporarily in the auxiliary memory. Every few seconds the switching mechanism connects this auxiliary memory to the computer. The stored information is transferred rapidly to the main computer's memory, and the computer spends a few hundredths or perhaps tenths of a second working on this problem, before moving on to the next. The computer services all terminals so rapidly and returns to any particular user so frequently that the operator at each terminal has the feeling that he is in constant communication with the computer. The chief advantages of time-sharing are that the computer's cost may be distributed among the several terminal stations, which may be installed in locations remote from the computer, wherever they are most convenient as, for example, at a charge-out desk, or in cataloger's office.

The computer can be, and has been, adapted to a variety of library tasks, by virtue of its ability to process and display information at high speeds, to perform repetitive tasks with precision, to carry out logical decisions which a programmer has prepared, and to provide multiple access to and from remote locations.

So much has been written and said about the computer, and its genii-like qualities, that it is well to be aware of its limitations. The first is that it does not truly "think" for itself, but carries out a pre-planned series of operations, albeit this may involve delegating to the computer certain logical "decisions." More importantly in the present context, the computer has a limited core memory. The largest computers have the capacity to store the information contained in a small library card catalog. None have anything approaching the size of memory required to store the contents of, say, a college library. Fortunately, there is really little reason for doing so. This point is effectively made by the Educational Facilities Laboratory.

> Data or factual evidence of small unit size is easily manipulated by the computer, but much more complex bodies of though or of knowledge are not. As of the present, it would appear that most of the literature in the humanities and social sciences will remain primarily useful in book form. There is no single advantage in converting Plato's dialogues to machine readable form and retrieving them or juxtaposing them by the computer. It is therefore economically senseless to attempt a massive conversion of existing library books to machine readable form, since the advantages of doing so are minimal in many subject fields, and the costs are enormous. The change to computer

> storage of full texts, when it comes, will be evolutionary, not revolutionary, and it is unlikely that the library as a repository of books will be replaced in the near future by a computer in the basement consulted by remote consoles. (25, p. 14)

Aside from the purely clerical chores of accounting, keeping inventory, maintaining circulation and statistical records, where the computer has amply demonstrated its value, the most promising area of library activity in which the computer can be effectively utilized seems to be in the area of bibliographic access. The most effective indexing terms for a given document change with time. As a discipline expands and the literature profiferates, it becomes necessary to increase the depth of indexing in order to hold the list of documents retrieved in a bibliographic search to a manageable minimum. It is much easier to add to or delete information in a computer memory than it is to change the index on a printed page or the information on a printed catalog card. Indeed, the computer can be programmed to add its own index entries to a given document in response to a user's request, without external intervention. Information can be inserted, in the middle of a sentence, if desired, or deleted, rearranged or modified in countless ways with lightning rapidity, once the record has been converted to machine-readable form.

To date, relatively little information exists in machine-readable form. The cost of converting even bibliographic records is generally too expensive for any one library to undertake. Estimates ranging from 50¢ to $1.50 have been reported for the cost of converting the bibliographic information contained on one library card.* The undertaking by the Library of Congress in its MARC project discussed more fully elsewhere, makes it feasible to consider computer processing of new material, but a broad national effort is still needed to convert the retrospective indexing information representing the bulk of most libraries' holdings.

---

*Lawrence Buckland, Conference on Implications of Modern Technology for Small College Libraries, Washington and Lee University, Lexington, Virginia, January 15, 16, 17, 1968.

At present costs, only the largest libraries can afford to undertake the development of computerized systems because of the large investment required for research and development. These experimental efforts are required to pave the way for other libraries, but substantial subsidies are required even for these efforts. Further standardization of machine languages and computer components will be necessary before these experimental programs can be made available on a broad basis. Package programs are under development by a number of libraries and commercial firms, and these will assist the introduction of computers in libraries; nevertheless it may require another five or ten years for these efforts to come to fruition.

It is likely that cost of equipment will be reduced in the future so that economics will become less of a consideration, but part of the economic advantage of sharing computer time with other organizations may be delayed by technical considerations. The present limit of 30 terminals linked to a computer will undoubtedly be extended, but another limiting factor is transmission capability. It is stated that:

> ...this requires a broad band transmission channel, the line costs are very high. The costs of microwave transmission are still higher. Even more restrictive is the lack, at present, of a national switching system that would allow rental of a broad band channel for a few minutes of time to any geographic location, which interlibrary transmission requires. And even if educational communication satellites were available today to facilitate transmission, the lack of an adequate switching system on a national scale would prevent facsimile transmission of text from becoming a common means of information exchange for libraries.(25, p. 14)

Further information regarding computer technology and its characteristics are contained in a comprehensive summary contained in Carter's report (13) on national library networks. Most of the figures contained in the foregoing analysis were derived from this source.

Microfilm Technology and Libraries

Microform is already a familiar storage medium in most libraries. It is the primary storage medium for newspaper files and other materials difficult to acquire in original format, for permanent storage of fragile, valuable or perishable documents such as archival manuscripts, and as a replacement for bulky volumes of little-used material. There are a vareity of forms available, the most common being the 35mm film in one hundred foot reels such as used for the New York Times. Recently microfilm has appeared in 16mm size. The film is usually advanced manually in a reader, one frame at a time; however, cartridge inserts and motorized units, to advance and rewind the film and to eliminate the necessity to thread the film onto the reader, are coming into use.

Another microform is the micro-card introduced by Fremont Rider. This is an opaque or positive print of a number of photographically reduced pages on a card, 2 x 5 inches in size, or 4 x 6, 6 x 9, etc. Bibliographic information in regular sized print is generally given at the top of the card to assist location of the position of the text required, but the text itself is read in a special reader.

The preferred form in governmental agencies for micro-materials is now the microfiche, a positive or negative sheet of film which may be of any size but is generally 4 x 6 inches.

Introduced initially as a means of reducing storage costs or preserving fragile material, microforms have tended to be an inexpensive means of acquiring scarce source materials such as rare books and manuscripts and government documents or of preserving bulky newspaper files. More recently it has become a publication medium for government reports such as those emanating from the Department of Defense, the National Aeronautics and Space Administration, the Office of Education, and the Department of Commerce. As a publication medium, microforms involve a separate and special technique together with precision equipment. They are prepared on a precision camera employing high-resolution film and then developed in a processor geared to very close limits to ensure quality. After inspection, the film may be reproduced in quantity, cut, packaged, and labeled for distribution. These processes are generally performed by commercial companies following careful specifications to ensure quality and the ability to make successive copies.

The technical problems of mass producing documents have been sufficiently resolved so that microforms are widely used in government and industry, and production has become a billion dollar industry in the United States. Over 100 sources of microforms and equipment were identified in 1967 (53) with additional sources in other foreign countries. University Microfilms, a major producer in the U.S., has a wide range out-of-print books, newspapers, and journals available on film and is the principal source of doctoral dissertations. The New York Public Library recently contracted with the 3M Company (Minnesota Mining and Manufacturing Co.) for the photographing and distributing the equivalent of over 100,000 volumes of materials.

One of the chief virtues of microform is its low cost compared to other forms of storage. Produced in quantity the page rate drops to a fraction of a cent per page, considerably below regular publishing, depending on the reduction ratio employed. Further savings are effected in the cost of storage resulting from the extreme compaction of information. This is dramatically stated by Leimkuhler in these terms:

> The technical limits to miniaturization have hardly been approached. Commercially available techniques can achieve reducations in the size of a manuscript of 500 to 1. (The federal-standard microfiche reduces a manuscript to about one four-hundredth of its original size). Experiments have achieved reducations of 1 million to 1, which would permit a library of 5 million volumes to be recorded in the space of 5 books. There has even been speculation about recording information at the molecular level and achieving reduction ratios of 10 billion to 1. This would raise the possibility (or threat) of recording all of man's knowledge on a few sheets of paper.
>
> High-ratio microeducation offers the prospect, not only of greatly increasing storage capacity, but also increasing the local self-sufficiency of libraries. Verner W. Clapp has calculated that,

> with microfilms of 200-diameter reduction, a library could add a million 250-page books for only $18,750... (34, p. 15)

Utilization of microforms in a library may entail special file equipment, microfilm readers, and devices to make a full-sized reproduction.

There are, of course, attendant costs in the use of microforms including special equipment for storage, reading, copying, and locating film. It is estimated that if a working collection of one million volumes was stored in ultra-microform serving an average university community, that the number of sets of films required plus the number of readers, copying devices and special handling equipment required would bring the per page cost back to a rate comparable to current printing.* The million volume collection on ultra-microform, originally proposed by David Hays at the Rand Corporation, has considerable appeal, however, because the production, indexing and processing costs of the package are incurred only once; and if enough libraries acquired it, these costs could be distributed over a broad base. The estimate is based on one thousand identical sets.

Two difficulties arise in a library containing large numbers of microforms and many users: file maintenance, and the queueing problem. These problems associated in government agencies with large microfilm files have led to a number of devices for automatic retrieval of documents from the file. Some of these devices operate with roll film as in the case of the File search system or Miracode. Others use fiche units as in the Minicard system. High speed retrieval units, such as the Verac by the AVCO corporation, have been developed experimentally; the Walnut system, developed by IBM for the Central Intelligence Agency, is an operating system. Many such retrieval devices are experimental models or designed for business or industrial needs, and have limited capacity in terms of library requirements. Although access time is available in several seconds, there is no operating experience to judge the volume of requests which could be handled. Present equipment of this type is expensive and approximates the cost of the data stored in it.

---

*Letter from Melville J. Ruggles, Program Officer, Council on Library Resources, Inc.

An additional feature of the document retrieval devices is the ability to combine the text of a document with the associated indexing. This feature permits rapid scanning of documents on a subject, similar to browsing through the stacks, and automatically adds additional aids for the selection of material. In some cases this is achieved by coding on the film, much like a sound track on a motion picture film. This approach would seem to be a disadvantage because of the difficulty of revising indexing terminology. For reasons already given, it seems more desirable to use the computer for bibliographic control. This limitation could be overcome by connecting a computer to the retrieval device. Computer indexing can be up-dated more readily and can be employed to locate documents. The approach offers the further advantage of monitoring the use of the document file to determine the amount of utilization documents receive, the subject approach used to locate specific documents, and the portions of documents which are actually used.

Other characteristics of microform storage and retrieval devices which extend their utility are the ability to reduce all documents to a uniform medium so that form divisions attendent with existing collections are eliminated. There are limitations in the ability to handle color, but presumably these will be resolved. The merging of forms also facilitates the merging of indexing information. The ability to scan the file and copy at will provides the ability to restructure and sequence information of all types much on the order of the custom text-book referred to earlier. This medium also facilitates purging and up-dating the file.

In spite of the techniques available for large scale production of microforms and the sophisticated devices being developed to store and retrieve them, this medium is used by libraries primarily for little-used materials. This situation is likely to continue for some time for a number of reasons. In the first place, copyright laws restrict the reporduction of materials, so that only a portion of the resources can be converted to microform. It is not economically feasible to convert books to microform unless there is a large demand for each unit produced. (15, pp. 5-15) Beyond this, the materials to be copied represent such a wide range of type fonts, paper quality or photogenic characteristics, diagrmas, charts, illustrations etc., that they have not been carefully analyzed to set appropriate standards for reduction ratios. Lack of standards in the industry gives rise to wide variety of types and sizes of microforms and the associated equipment necessary to use them. In general, different readers are required for roll film, opaque microcards, and for microfiche. The quality of these readers is such that there is strong user resistance to the medium. In the opinion of William R. Hawken, one of the leading microform experts in the United States, there are no suitable microform readers available today, and the cost of the equipment available is too high for general use. Microfiche appears to show more promise than roll film or opaque cards because the problem of optics is simpler and reading equipment more reasonably priced. He states that unitl microforms can meet the minimum needs of functional equivalence to the book, their utility in libraries will continue to remain limited, the market, as viewed by the manufacturer will continue to remain small and unimportant, and the hope and promise that scholars and librarians alike have clung to for thirty years will continue to remain unfulfilled and unrealized.

The reader has little option in using microforms than the present limited reading equipment, for there is no suitable copying device that will restore the print to original size at a reasonable cost. The quality of such devices is generally considered to be poor, but a number of new products have recently been introduced which have not been fully assessed. Even if the quality of these devices were improved, the page cost to the user would still be relatively high (circa 10¢ per page).

The overall assessment of the microform medium at this time is stated by the Educational Facilities Laboratory in these terms:

> At present, the usefulness of microform technology is minimized by the fragmented nature of the industry. This has resulted in a lack of standardization in the production of microforms. The past few years have seen a rapid multiplication in the kinds and shapes of microforms which are not compatible, have little or no relationship to each other, and are in fact mutually exclusive. Advances in use will depend on standardization of a reasonably restricted number of kinds of microforms, and on the production of a system of machines

> interrelated so that their output can be automatically converted, quickly and cheaply, from one form to another. It must be possible to convert from micro to micro, micro to large and large to large, producing retention copies of such quality that copies of the copies can be made with no deterioration of the image. In addition, and most important, there must be equipment for consulting microforms that is inexpensive, easy to use, and that has excellent legibility.
>
> These developments do not seem likely to occur in the near future. Progress in the field is slow since there is no single firm in the industry that plays a dominant role in setting the pace for all of the others...(25, p. 13)

In summary, the development of new media for the storage of index and textual information has attracted considerable interest and offers many attractive features for the long term solution of library problems. As the most glamorous area of speculation, it has attracted numerous science fiction accounts of the library world of tomorrow. At present, however, we have only the experience of a few libraries on which to generalize. The National Library of Medicine in the MEDLARS project, the most conspicuous example of computer indexing, prepares a comprehensive index of current medical literature by computer technology and produces the Index Medicus in printed form or for computer search on magnetic tapes. This project has been universally acclaimed as a genuine technical advance. With respect to microforms, there exist in several governmental agencies, such as NASA, the Department of Defense, and the Department of Commerce and in industrial organizations, such as IBM, large files of reports in microfiche. These organizations have demonstrated the utility of this medium for compact storage and inexpensive duplication and distribution of microfiche copies, or if desired full-sized reproductions. These methods are being extensively tested in academic library settings by MIT in its project Intrex.

It cannot be assumed that current experience in these governmental agencies can be transferred automatically to all libraries or even academic libraries in its present form, for there are a number of areas which require more careful investigation. A typical academic library represents a wide range in format of material covering centuries of publication. It cannot be assumed that these can readily be transformed into uniform microfiche as in the case of government reports, which are prepared according to standard specifications. Libraries have no direct control over the publication of materials and would need to impose standards after the appropriate specifications have been established. Very little information exists in either machine-readable form or microform, and most of what does exist varies appreciably in format and is incompatible. Project MARC at the Library of Congress will effectively stabilize bibliographic information in machine-readable form, but this will cover only current American imprints. Numerous other types of bibliographic information need to be brought under control such as periodical indexes, document indexes, and report literature. The most conspicuous element lacking in the entire picture is the degree of coordination required between libraries to attain definition of requirements through careful investigation and the implementation of standards once they have been achieved.

From the standpoint of the individual college library there does not appear to be any immediate action which can be undertaken to improve existing conditions. New media for the storage of both textual and bibliographic information will eventually have an impact on the operation of college libraries, but they are in no position individually to alter current conditions.

Within the limited sphere of factors which individual libraries can control directly, the application of technology, according to the definition cited earlier, requires employment of analytical techniques as well as the consideration of available alternatives. The analysis step, called systems analysis, is a detailed description of existing procedures. This step provides a frame of reference for consideration of other alternatives and is essential to any study of efficiency. The latter procedure, referred to as systems design, is a careful, objective method of assessing various approaches to accomplishment of an objective under consideration. Technology does not imply that a process will be mechanized, and the best solution may be redesigning a manual operation.

The application of machines normally requires a high volume of repetitive operations which can be standardized and which will result in lower unit costs of work performed. These conditions generally obtain only in the largest libraries and are economically unfeasible for individual college libraries.

Under current circumstances, therefore, the college library can gain some efficiency in applying the techniques available in systmes analysis and desing. It is unlikely that technical devices such as computers or microtechnology will provide significant improvement over current procedures. The limitations in application of technology faced by the average college library are: 1) more expertise is required than is available on the staff, 2) a relatively large expenditure is required to exploit technology, 3) the return on investment is likely to be small in view of the small size of most college libraries.

Although technology currently offers little relief for college libraries, acting independently, it does suggest approaches for the improvement of library.service. The basic question is whether sufficient volume of work can be developed through cooperative arrangements and whether unit costs can be significantly reduced in the process.

Chapter VII

Summary and Recommendations

The history of American higher education, from its inception, reveals a continued expansion of programs, a diversification of instructional techniques and an up-grading of quality. This process is accelerating today. College enrollments have risen sharply in this century, and the projected growth pattern indicates that enrollments will more than double in the next two decades. The rapid growth of knowledge and attendant technologies, coupled with enrollment pressures, force a more careful assessment of the role of the college and the appropriate means to achieve its objectives. The implication for libraries are tremendous.

The library has emerged as a vital part of the instructional process because it is the only means to deal with knowledge in a broad and systematic fashion. Classroom instruction has been forced to be more selective in transmission of knowledge, and learning appears to be enhanced when greater responsibility is placed upon the student for acquisition of information. The primary objective of instruction appears to be the transmission of the present status of the various disciplines. The instructor serves as guide and critic while the student explores the conceptual framework of the discipline, the nature of evidence and methods employed to establish it, and engages in some independent exploration.

The average college library resources are too deficient and the rate of growth too low to provide the resources required for quality instruction. Without increased support or more efficient access to information, instruction is forced to compensate by oral transmission of a more restricted body of information.

The plight of the college libraries are, in a sense, common to most academic libraries, but the economic constraints are more limiting. The thrust of most of the efforts to alleviate library problems is directed to large research libraries. The college library will benefit only indirectly unless more deliberate attention is directed to its needs.

The impact of current technological developments would not appear to offer much relief for college libraries, acting independently, in the immediate future. The more promising devices are geared to a higher volume of work than that represented in the average college library. But the application of technology (being conceptual as well as mechanical) would still call for a systems analysis approach to determine what applications can and cannot be used. Further, there are two avenues which might be employed to exploit mass production techniques: increased standardization and greater coordination of effort. The present circumstances are well described by Patterson and Longsworth in these terms:

> At present, higher education is largely an assortment of discrete and varying components. Each is isolated from the other for the most part. In isolation, each pursues its own program with whatever resources it has. In the process, wasteful duplication of effort is common; so too are shortages of resources in all the common fields.
>
> What appears needed instead is not the absorption of all component institutions into a closed system or monolithic organic whole. The greatest actual need, for the welfare of the whole educational enterprise, is to establish and radically strengthen the cooperative association of diverse institutions in regions or localities, or in special areas of interest. (46)

One avenue of approach is to define the resource requirements of the college library more rigorously to insure that the prime resources are immediately available and that there is a formal means of access to supplementary needs. The concept of the core library seems most useful in this respect. Existing tools are limited in their ability to identify the core. Indeed, little has been done to establish empirically the existence of the core despite the fact that most libraries and faculty members agree that one exists. The epistemological basis for the core is in the observation that knowledge itself tends to abstract and cumulate on the basis of concensus. The instructional process tends to be even more selective.

Further attention to this phenomenon would demonstrate that there is a core of literature, defined by concensus, that represents the findings of a discipline. This concensus will be found by observing three basic kinds of material in a discipline: observations, theorizing and synthesis. There is some practical basis for believing that there is a core based upon several studies showing overlap in even uncoordinated acquisitions among groups of libraries. There are also evidences in the current use of standard lists such as Books for College Libraries. If a core is identified then each college could have a body of source material, while less intensively used material could be pooled in a common store serving a number of institutions.

The core if used would be defined by levels and by subject areas, provide for deletions, speed up selection process and tend to keep books in print that are most useful.

The most important advantage to the selection of a core is the potential for savings in money and skilled manpower which can be diverted to improved services to students. The present "go it alone" approach is unacceptable in meeting today's needs in selecting, ordering, cataloging and processing books. Some attempts have been made to effect savings through groups of colleges entering into an arrangement characterized in this report as Centralized Bibliographic Computer Store. This has several advantages but little evidence of real savings.

Processing centers which order, catalog and process books for a group of libraries show some evidence of savings. Such centers now exist through both public and commercial groups. However, the greatest saving could be effected through adding to the actions of the processing centers automatic purchasing. The savings thus effected make it essential to devote maximum efforts to defining the core.

While the establishment of the core and full utilization of process centers are the major problems, there are other areas of technological advance to be considered. The improvement of bibliographic access to library resources is basically a national problem beyond the scope of the individual college library. Numerous intellectual and organizational problems are required to effect any basic changes. The most immediate improvement anticipated is machine-readable cataloging information (MARC tapes) from the Library of Congress. This is a significant step toward providing an avenue for mechanized bibliographic searching. A further impact of the MARC tapes is that they facilitate the automation of library processing. It is not anticipated that it would be practical for the average college library to undertake automation as an independent effort because of the limitation of equipment and technical staff. A more practical approach would appear to be a cooperative arrangement with a group of libraries in a region or to sub-contract the operation to a processing firm. An important by-product of library automation will be wider access to information about specific holdings in other libraries.

The physical access to information through interlibrary loans may be improved through technology. Increased knowledge of available resources through printed catalogs and union catalogs prepared from machine-readable cataloging information will be a major assistance. Group storage should assist college libraries to have a greater total of documents from which to draw.

Microforms appear to be a promising approach to increased resources at lower costs, but the lack of standards in the industry and the lack of adequate viewing and copying equipment will retard developments in this area. This format shows promise of meeting needs for little used materials but not under present circumstances, as a general utility medium for textual material.

While it is true that the motivation for the development of information systems stem primarily from the science-technology information and document problem, the ultimate goal is an information network encompassing the information requirements of all users.

It is clear that such a concept is beyond the range of immediate feasibility. Elements of the plan, however, are feasible on a more limited geographical basis using slower communication devices initially. A basic ingredient of the concept is the establishment of a coordinating agency to formulate plans and procedures, to develop standards, and to perform specialized services to minimize duplication of effort. This approach appears feasible on a regional basis, and some elements seem desirable on a national level. The primary problem is to develop a large enough volume of work to be able to employ current technology advantageously and to exploit as much mass market potential as is consistent with local requirements. Some of the elements envisioned in the information network of the future are:

1 - a local library service unit containing a basic collection tailored to the needs of its clientele with a communication link to a regional center.

2 - regional centers to serve as a backstop collection of broader scope and of lesser used materials with specialized bibliographic and reference services to supplement the local library and a communication link to national centers.

3 - national information centers containing comprehensive resources and services to satisfy all known user requirements.

4 - control of document production through standards to insure uniform format, stylized title page and abstracts to facilitate analysis and indexing (original preferably in machine readable form).

5 - sophisticated analysis and editing in central agencies to evaluate significance and possible utility of documents.

6 - facility for automatic selection of documents through specifications of needs and interests of local libraries.

7 - centralized acquisition and processing of documents for local and regional libraries.

8 - regional computer facilities to assist in local record keeping and in administrative functions and to monitor use patterns.

The maximum benefit to such group actions, and the primary recommendation of this report is the package library, built on the core of the disciplines, and common to all colleges offering the same program. Such a package library, purchased and processed together with all additional books through a processing center, offers economic advantages to participating colleges and the advantages increase as the number of participating libraries increase.

It is evident that if college library services are to attain their potential and if technology is to be employed advantageously, more specific attention needs to be directed to college library problems. The most expeditious approach would appear to be a coordinating agency on the national level to provide specialized assistance and to coordinate research in this field. Among the special services required are:

1) establishment of a comprehensive editorial service to assist in the definition of appropriate resources of all types for a basic collection, current additions, and deletions. The concept is an expansion of Choice activities employing the pooled judgments of subject experts and the monitoring of actual use to identify more accurately, on a continuing basis, the prime resources.

2) guidance and assistance in the establishment of regional back-up collections for pooling little used materials and establishing machinery to guarantee access to these materials.

3) formulation of standards for processing materials to enhance economical use of centralized services.

4) formulation of standards on equipment requirements based on assessment of needs to exert maximum influence on equipment manufacturers.

5) development of specifications for, and provide advisory services to assist in, the establishment of centralized services.

6) exploration of methods of financial assistance to insure acquisition of a basic collection in each college library.

The objective of the centralized agency would be to insure as much coordination and standardization as possible consistent with individual requirements. Defining common needs would influence the market to meet requirements at economical prices, and to encourage needed products through identification of the potential market. These services require supporting research and development activities aimed at college library problems and coordinated with similar research endeavors. Among the projects required are:

1) evaluation of the core collection concept.

2) analysis of use of college library resources to improve specifications of needs, to refine up-dating collections, and to control the size of campus libraries.

3) identification of requisite professional skills to guide library school instruction.

4) testing of various approaches to bibliographic and physical access to determine appropriateness in a college environment.

5) development of economic studies to identify optimum patterns of organization, equipment specifications, etc.

The network concept for the college library does not relieve the college library of the responsibility to maintain an active, working collection of the most used resources, but it does make it possible to define these resources more rigorously. Guaranteed access to lesser used materials through centralization of common functions and access to specialized services by groups of libraries would enable the individual library to concentrate attention on the most pressing needs of local users.

Considering the current status of technology and library planning, college libraries should, therefore, explore new approaches such as:

1) find more effective techniques to define the resources which are needed locally.

2) identify the lesser used materials which might be held jointly with other libraries.

3) explore automatic approaches to the acquisition of material.

4) standardize bibliographic and processing routines.

5) explore methods of sharing processing costs.

6) develop more effective bibliographic and physical access to local and shared resources.

7) develop appropriate coordination mechanisms to achieve the above.

The college library would appear to be a fruitful area for research in and testing of, new concepts and technology. Not only does the college library constitute an area of pressing needs and limited financial ability, but it represents a smaller, controlled environment for investigation. There would appear to be greater likelihood that a core collection amenable to automation techniques can be identified in college libraries than in the larger research libraries. More complete bibliographic control in this area would not only enhance the development of collections and the sharing of resources but success in this area should be broadly applicable to other libraries as well.

Appendix I

Theoretical Basis For a Core Collection

It is generally asserted that inquiry begins with a problem encountered in experience which produces tension and uncertainty. (44, pp. 16-17) Empirical research is interpreted as a form of adjustment technique analogous to earlier devices such as magic, animism, mythology, theology, and metaphysics. (33) Such empirical research emerged during the Sixteenth Century from the merging of mathematical and empirical processes and has demonstrated its superiority as an adjustment technique through its ability to produce reliable information. Inquiry, however, is two-dimensional. It is directed toward the solution of an immediate problem and an explanation of the problem. Many problems have been resolved through trial and error before an adequate explanation was discovered. Primitive people learned to adapt themselves to their environment and develop technology with varying degrees of success by trial and error. Many of the technological changes occurring during the Industrial Revolution evolved in the same manner. It seems that initially explanation serves to relieve tension and anxiety in areas where trial and error has failed and uncertainty is great such as the application of magic by primitive people in situations where adjustment is incomplete. (41) But empirical research in modern times appears to be more than a superior method of acquiring adequate explanation for puzzling phenomena, for it is systematically applied to the realm of technology as a means of rationalization and improvement.

Empirical research is considered to be a refinement of the basic processes which are employed in common sense adjustment to problems encountered in daily life. "In our historic era empirical sicence criticizes, augments, and systematizes practical experience. The science of one generation become incorporated in the technology of the succeeding one. Science and practice cooperate in the adjustment of man to his environment." (36, p. 1)

The distinction between science and common sense is essentially a difference in objectives and in the rigor of the reasoning applied in the interpretation of experience. (18, pp. 42-62) Common sense is primarily concerned with the solution of practical problems, whereas science is concerned with hypothetical, universal situations and the formulation of systematic interpretations. This distinction tends to exclude much knowledge commonly denoted as scientific because of its esoteric nature and the highly refined methods by which it is acquired. It emphasizes the theoretical and systematic approach to the development of a body of information for its own sake. The difference between "pure" and "applied" research is essentially that of a general theoretical objective in the case of the former as contrasted to an immediate, practical objective in the latter case.

A further characteristic of empirical research is that although a problem in science may originate in practical experience, it ultimately evolves from scientific activity. Werkmeister states that.

> However, in consequence of centuries of scientific tradition, the scientist no longer waits for the problems to arise, in haphazard fashion; he looks for them and forces them into prominence through experimental disturbances of the 'normal'course of events. Many of his problems are the result of his own theories and are suggested to him by his own imagination. Others he creates by deliberately modifying the conditions under which given phenomena occur. Still others are entailed by solutions of problems thus found or created. (55, p. 299)

The specific objectives of empirical research are expressed in various ways. The statement which occurs most frequently is that empirical research attempts "to describe particular phenomena in the world of our experience and to establish general principles by means of which they can be explained and predicted." (28, p. 1) A more ambitious claim which is perhaps typical of those expressed in philosophy states that:

> ...science, starting with adequate and concise descriptions of individual facts and events, aims at an explanation of these facts and events through recourse to universal and necessary laws, and that it sees its ultimate aim in the unification of all laws in one explanatory system which integrates the whole of experience and makes it deducible from the basic postulates or principles of the system. (55, p. 335)

A further view expresses the objective of empirical research in a form which may be considered more realistic in terms of the immediate operations employed by an investigator:

> Science is an interconnected series of concepts and conceptual schemes that have developed as a result of experimentation and observation and are fruitful of further experimentation and observation...Science is a speculative enterprise. The validity of a new idea and the significance of a new experimental finding are to be measured by the consequences--consequences in terms of other ideas and other experiments. Thus conceived, science is not a quest for certainty; it is rather a quest which is successfully only to the degree that it is continuous. (18, pp. 25-26)

The latter definition of objectives is a cautious statement which deliberately avoids involvement with metaphysical problems concerning reality. All of the statements cited emphasize the quest for integration of experience and differ essentially on the significance of the product. To be more specific on the latter point necessitates either various metaphysical assumptions or consideration of the goals in utilitarian terms.

Designation of the subject matter of empirical research is expressed in terms of high abstraction. Objects, events, entities, or phenomena are the manner in which the substance of research is indicated. Some consider phenomena or events as the ultimate substance of experience and objects as the interacting elements producing events. An alternative interpretation is stated as follows:

> Recent literature in the philosophy of science seems to have adopted the term event as the most adequate name for the basic entity of science...almost everything that one may mention is an event -- durations, spaces, relations, and numbers, things as well as processes, happenings in the physical realm as well as in the 'minds'. The realm of the known may therefore be described as a complex of events. The events bear to one another relationships of various kinds - - spatial, temporal, causal, similarity and difference, equality and inequality of magnitude. These relation-

> ships determine associations and dissociations of events, containing and contained events, dependent and independent events, elemental events, collections, organic wholes, and Gestalts. (4, p. 50)

Another alternative expression which appears to predominate in the current literature states that:

> The initial objects of science are the things experienced in perception, and their most general character are positions in space and time. The systematic and ultimately quantitative investigation of space-time order may be called generalized physics. Thus one arrives at the doctrine of physicalism, which asserts that the concepts of empirical science are reducible to those which express the properties of spatio-temporal things. (36, p. 3)

Such statements do not fully clarify the substance with which empirical research deals because there are a number of dimensions involved. Spatio-temporal things may be considered as whole or they may be reduced by analysis to component parts. Spatio-temporal things may be identified by various characteristics. Some of these characteristics may be stated in terms of physical properties such as color, shape, or texture, and other characteristics may be identified in terms of how a thing responds to other things under certain circumstances. Symbols are employed to express the unit of the thing being examined or its properties and relations and is designated as a concept. The latter constitutes any specific content of experience which can be distinguished and is meaningful. Only the definition of these symbols distinguish between the various dimensions of the substance under observation. Although the term event can be defined in such a way as to encompass all of these distinctions, for purposes of clarity there is some merit in using the term object to designate the spatio-temporal things of perception and the term phenomena to designate the actions of things in relation to other things.

The task of empirical research is considered, then, as a process of formulating principles to provide a means of prediction and control of the objects of experience. The formulation of principles is the ultimate goal, but it is preceded by a series of operations directed toward increasingly precise definition of the objects of experience and determination of their interrelationships. The procedure is one of successive definition and successive integration. The latter constitutes the scientific explanation, but it is not a final explanation. It serves to interpret how the substance of experience functions but does not answer the question of why it functions as observed.

In the initial stage of empirical research a specific problem encountered in experience is analyzed for purposes of identifying the elements or factors which appear to be relevant to its explanation. Whereas common sense is content with manipulation of these factors or the introduction of new elements by trial and error to find a workable solution, empirical science probes deeper to determine the basis for the manner in which the factors interact. The latter course leads to the examination of other similar cases and the more precise identification of these common elements. The latter can then be abstracted and defined in terms of distinguishing characteristics for further examination. This process of abstraction results in the creation of a concept. The definition of the concept is achieved by selection of the characteristics or attributes which appear to be essential among its manifold characteristics in a particular context of experience. The concept is initially a rough approximation designated by attributes which may later prove to be irrelevant. Continuous observation by a society of observers ultimately produces a more precise identification of essential attributes through a process of elimination. Lazarsfeld describes concept formation in the social sciences in these terms:

> No science deals with its objects of study in their full concreteness. It selects certain of their properties and attempts to establish relations among them. The finding of such laws is the ultimate goal of all scientific inquiries. But in the social sciences the singling out of relevant properties is in itself a major problem. No standard terminology has yet been developed for this task. The properties are sometimes called aspects or attributes, and often the term 'variable' is borrowed from mathematics as the most general category. The attribution of properties is interchangeably called description, classification, or measurement, using this last term in a very broad sense...

> The first step seems to be the creation of a rather vague image or construct that results from the author's immersion in all the detail of a theoretical problem. The creative act may begin with the perception of many disparate phenomena as having some underlying characteristic in common. Or the author may have observed certain regularities and is trying to account for them. In any case, the concept, when first created, is some vaguely conceived entity that makes the observed relations meaningful. (33, p. 15)

Ultimately the concepts of science tend to relate to entities which cannot be perceived directly. They are known only in terms of their effect on perceptual objects and are sometimes designated as a "construct" to distinguish them from concepts derived from objects of perception. It is asserted that:

> The field of natural science is initially constituted by perceptible phenomena, but explanation extends the realm of nature by positing objects that are not perceptible in the usual sense of the term. A phenomenon may be explained as the effect of action by a body that later may be directly perceived. This procedure is exemplified by the perception of Neptune after its position had been predicted the the theory of perturbation. Vitamins were initially posited by hypothesis for the explanation of physiological phenomena and were then isolated. Similarly, genes were instrumental for explaining the laws of heredity and were then located in structures perceived through a microscope. In building microphysical theories, however, objects are posited that according to physical theory can be known only through their effects. The visual datum of a natural thing is ascribed to radiation that it emits or reflects. Scintillations of an appropriate screen, condensation tracts in a cloud chamber, and indications of Geiger counters are attributed to the ionizing action of electrified particles. The only means of cognizing such hypothetical objects is through perceptible phenomena that they serve to explain. (35, pp. 114-115)

The achievement of precision in the definition of concepts is accomplished in a number of ways. The initial product of perception is a qualitative description of various attributes of an object. Through repeated observation employing more refined procedures some attributes can be eliminated from further consideration and other attributes may be discovered. Sometimes this is accomplished by extending the range of observation through new tools such as the microscope or by controlling the observational conditions more rigorously. Various quantitative devices may be introduced to lend further precision to the selection and comparison of objects. Much of what is called the methodology of research is concerned with precision in observation and measurement. A number of practices have become established conventions, but techniques are continually being developed to meet new problems. Physical quantities such as counting, measurement of length, time, weight, etc., in the physical sciences are defined by means of experimentally controlled procedures which become standardized practices. (35, p. 112) This method of determining properties has been designated by P.W. Bridgman as an operational definition. (9) Werkmeister describes the development of precision in these terms:

> Where perceptual objects reveal a manifoldness of gradual transitions and 'fluid' contours, the scientist insists upon concepts which provide sharp delimitations, and which eliminate all vagueness of interpretation. Instead of remaining satisfied with the indefinite distinctions of 'more' or 'less,' 'nearer' or 'farther,' 'stronger' or 'weaker' -- so characteristic of perceptual experience -- he introduces a scale of numerical values by means of which he succeeds in transforming experienced warmth into the concept of 'temperature,' 'kinesthetic' sensations into the concepts of 'force' and 'weight,' and, in general, any experienced content into some objectively conceived and quantitatively determinable concept. The concepts of science are, therefore, in a specific sense a step removed from the concepts of perceptual experience. The 'data' of science are no longer our perceptions as such, but our perceptions quantitatively fixed and transformed. (55, p. 335)

Often the product of investigation in the early stages of research does not permit formulation of a scientific explanation. As Lenzen states:

> Scientific inquiry is initially directed to the analysis and description of natural things in the status of perceptable objects. For example, if a specimen of a new kind of mineral is discovered, the task of scientific description is to record its place of occurrence, to specify its properties, and to name it. The mineral is first characterized by superficial qualities -- for example, gold is described as yellow, lustrous and heavy ... Descriptions of perceptable objects with the aid of concepts furnish the data for the systems of classification of the descriptive natural sciences. (36, p. 110)

Classification, however, is only a first approximation of the relationship between such entities and may be quite arbitrary. Scientific explanation requires the determination of a particular type of attribute -- the functional characteristics of objects. Werkmeister states that:

> ...the real 'essence' of a thing lies not in some definable substance, but in the relations the thing has to other things, and in the way in which it reacts to these other things; it lies not in the sense qualities disclosed in first-person experience, but in the measurable properties -- such as 'exponent of refraction' or its 'magnetic susceptibility' -- which place it at a certain point in a scale and which are derived from a quantitative interpolation of observations. These measurable properties are the basic constants in terms of which the scientist integrates experience, and the interrelationships of which he tries to express in his general laws. (55, pp. 335-336)

Observation of objects in empirical research ultimately leads, therefore, to a search for attributes which will reveal the functional relationships between objects and their precise measurement. It is asserted that:

> ...the further the scientific cognition of nature progresses, the more the concept of things are re-

> placed by concepts of relations. The concept of substance resolves into the concept of function. But the concept of function is not considered as the expression of a mere 'togetherness' of experienced facts, but as a genuine achievement of thought. It is the original form of connection as such, and of experience itself. The concept of substance is to be replaced ultimately by that of invariance.*

The ultimate goal of science, then is a statement of uniformities in objects of their invariant relationships. This objective is based ultimately on the premise that objects and phenomena of experience conform to a coherent and continuous pattern of inter-dependency.

As in the definition of thephysical properties of an object, the attributes describing relations are defined by successive approximations until the most significant relationships have been determined and defined with precision. Initially some of the relations of objects with one another may be discovered fortuitously. Hempel states that:

> The vocabulary of everyday discourse, which science has to use at least initially, does permit the statement of generalizations, such as that any unsupported body will fall to the ground; that wood floats on water but that any metal sinks in it; that all crows are black; that men are more intellectual than women; etc. But such generalizations in every day terms tend to have various shortcomings; 1) their constituent terms will often lack precision and uniformity of usage (as in the case of 'unsupported body,' 'intellectual' etc.), and, as a consequence, the resulting statement will have no clear and precise meaning; 2) some of the generalizations are of very limited scope (as, for example, the statement dealing only with crows) and thus have small predictive and explanatory power (compare in this respect the generalization of Archimedes' principle); 3) general principles couched in everyday terms usually have 'exceptions,' as is clearly illustrated by our examples. (28, pp. 20-21)

---

*"Substance," Encyclopedia Britannica, 14th Ed. XXI:497.

As a consequence of the limitations of generalizations stated in this form, empirical research attempts to formulate more precise, comprehensive, and reliable statements. The discovery of relationships which can be stated in this form is more than a matter of increased sophistication in research techniques. All the conditions implicit in a standard scientific experiment -- the rigid control of observational conditions, the isolation of conflicting factors, and refined measuring devices -- assist in the systematic elimination of concepts which do not reveal relationships and the precise measurements of concepts which have such capacity. But a significant statement of relationships is achieved by the abstraction of an element common to a number of statements concerning perceptual objects. This step is the product of imagination and ingenuity; there are no prescribed methods to follow. The process is described in a number of ways, such as postulation, induction, generalization, or abstraction and is illustrated by the following example:

> Thus, nearly all empirically observed behavior of bodies from the point of view of their movements in space and time are 'covered' by the general 'principles' of physics. That is, events as 'different' (from some points of view) as a man falling from a twentieth story window, a bullet fired into the air from a rifle, or drops of water in a rain storm, are all 'explained' by the same basic principle.
>
> No two cases of any of these events are ever identical in all respects nor are the natural conditions under which they occur ever the same. Yet by a process of ignoring all this variety and concentrating our attention on some single characteristic or aspect of the event (abstracting), we can make general statements that are equally true for all falling men, all rain drops, and all projectiles. (38, p. 6)

The statement of relationships is designated as a law and is achieved by abstraction from a group of seemingly disparate observations about objects. A law does not constitute a final statement of relationships, for it may be reformulated with greater precision or may be subsumed by a more general law. Werkmeister (55, pp. 264-265) cites an instance in the physical sciences of three levels of integration in the formulation of a law: 1) the functional interrelation of directly observable or measurable magnitudes (e.g. Galileo's law of falling bodies and Kepler's laws of planetary motion), 2) Newton's law of gravitation combining Galileo's and Kepler's laws, and 3) Einstein's principle of relativity integrating the laws of mechanics and electrodynamics.

It is a common misconception that empirical research consists in the discovery of facts guided primarily by curiosity. Parson states that:

> There is moe often implicit, than explicit, a deep-rooted view that the progress of scientific knowledge consists essentially in the cumulative piling up of 'discoveries' of 'fact'. Knowledge is held to be an entirely quantitative affair. The one important thing is to have observed what had not been observed before. (45, p. 6)

The predictive power of a science, according to this view, would result from the addition of discrete bits of information. But the realm of nature is too vast and too complex for any meaningful interpretation of the objects and phenomena in experience to be developed unless a more selective approach is employed. Observation is the basic method of empirical research, but it is not a random activity.

Observation is guided and directed to the acquisition of information which is significant for some purpose by successively refined devices. Initially, the problem prompting inquiry focuses attention upon elements which appear to be relevant to its solution and explanation. Identification of the common elements in related problems leads to the formulation of an approximate explanation. The latter corresponds to what is called a point of view or frame of reference. It may be entirely implicit at first but gradually is formulated in explicit terms. Its chief characteristic is the identification of the concepts which somehow must be taken into account in securing an adequate explanation. The concepts, as discussed previously, are specific contents of experience which are distinguishable and meaningful and are defined by means of attributes. Working within this general framework which prescribes the objects and phenomena to be observed, empirical research seeks to establish precise definitions of properties and relationships. The facts derived in this process consist of "empirically verifiable statements about phenomena in terms of a conceptual scheme." (45, p. 41) Such statements are designated in logica as propositions. A systematized body of propositions concerning the properties and relations of a given group of objects and phenomena is designated as a theory. It is a provisional explanation of one segment of experience. Its full elaboration would constitute a scientific explanation; but since integration of all the facts is not complete, it has only a tentative status. A theory represents a summarization of the knowledge attained by a discipline at a mature stage of development and is the basis from which further observation is made.

After a body of propositions have been attained by a discipline, systematic formulation can be accomplished by logical processes and is designated as axiomatic theory. Its chief components are a set of terms (i.e. concepts) and a body of propositions expressing empirically derived relations between them. The procedure for arranging these components by logical processes is illustrated by Zetterberg. (57, pp. 533-540) Theory, then, whether expressed in explicit form or employed implicitly in the derivation of problems, guides observation at a more advanced stage of research by indicating areas where knowledge is lacking or inadequate.

The chief probing tool of empirical research is the hypothesis. In simplest terms it is defined as an "intelligent guess" (57) concerning the facts to be found. Its primary function is to state explicitly the general objective of an investigation. In the latter sense it constitutes a "working hypothesis" (47) for a specific investigation, but the term hypothesis is frequently used interchangeably with theory. Terminology is inadequate to draw an adequate distinction between the two terms. Hypothesis may represent speculation or anticipation of facts and as such is the most refined proscription employed in observation. The conflict between hypothesis and theory arises when the investigator leaves the perceptual realm to postulate an entity which cannot be observed directly, for at this level of abstraction a single postulated entity may entail an extensive series of observations to establish confirmation. The chain of argument linking observational data with the central hypothesis and the network of minor "working hypotheses" derived from the central hypothesis, stated explicitly, may be similar in form to axiomatic theory. The only difference is that the former is entirely speculative while the latter has some degree of confirmation. For lack of more precise terms, it must be understood either that hypotheses vary in scope (in terms of the number of observations required for confirmation) and are speculative or that theories may refer to a systematic summary of investigations as well as a broad speculation scheme. The former definition is preferred in this investigation. The term hypothesis is used as an explicit statement of anticipated findings in an investigation as a "working hypothesis" or may entail numerous derived "working hypotheses." Theory is understood to mean a systematized body of empirically derived fact organized by logical methods.

The term theroy in the speculative sense is frequently used to describe broad "philosophical" speculation. This type of speculation appears to be characteristic of a discipline in its early stages of development. It represents a comprehensive explanatory system based on common sense views and imaginative insight but has little direct application in empirical research. It may suggest ideas for observation, but as it stands it has no empirical foundation. To be fruitful, speculation must take into account established facts and posit questions which are within the range of existing data and analytical equipment. The conceptual framework in a scientific sense evolves by stages from the facts established through observation and verification according to accepted scientific procedures. Imagination and speculation are channeled to the acquisition of facts useful in the extension of existing theory. Facts acquired in random fashion do not necessarily add to explanation in a scientific sense, for there are countless facts in any discipline which have not been incorporated into theory because their significance has not been discovered. The criterion of significance for facts is based on the extent to which they can be incorporated into the existing conceptual framework.

Ideally, then, observation in empirical research is directed in such a manner that the facts obtained can be cumulated and lead to the gradual unfolding of an explanation. The existing devices for this purpose are initially the frame of reference which prescribes the objects and phenomena for observation and ultimately a theory stated in a from amenable to rigorous logical analysis.

The hypothesis assumes importance as a tool in observation but is employed where facts are needed to extend the conceptual framework. The validity of the cognitive processes appears to be established by: 1) their occurrence in all the disciplines which have been analyzed, 2) the number of instances in which investigators working independently upon the same problem have reached identical conclusions, 3) the extent to which objectivity prevails in scientific endeavor without organizational direction.

The sociology of knowledge tends to view science as a cultural phenomenon and seeks to determine, among other things, how it comes about and is propagated. Whereas philosophy and science tend to regard the objectives and cognitive devices as constants, the sociology of knowledge tens to regard societal values as determinant. Science and society tend to influence and be influenced by each other; the nature and degree of this interaction is as yet unknown. Knowledge as interpreted by sociologists has been broadly defined and includes "...virtually the entire gamut of cultural products (ideas, ideologies, juristic and ethical beliefs, philosophy, science, technology: (42, p. 366) The available evidence seems to indicate that scientific knowledge is not a matter of immanent development. The scientist does not roam the realm of nature freely in search of truth; rather his attitudes, interests, and values are conditioned by various environmental influences, and the social expression of knowledge is encouraged or restrained by various political, economic, and institutional factors. It is asserted that:

> ...The concept of multiple causation is especially congenial to the academician, who has relative security, is loyal to the status quo from which he derives dignity and sustenance, who leans toward conciliation and sees something valuable in all viewpoints, thus tending toward a taxonomy which enables him to avoid taking sides by stressing the multiplicity of factors and the complexity of problems. (42, pp. 402-403)

Schroedinger attributes much of the dynamic concept in science in our era to the concept of evolution and change and states further that the popularity of a theory is science tends to influence selection of problems. The emphasis in the physical sciences can be attributed to various political and economic interests. Military interests have always stimulated science and technology. This influence is especially marked in totalitarian governments and in times of war. Newton's law of gravitation evolved from his preoccupation with researches in navigation and was not entirely a theoretical concern. (5, p. 313) Znancieki (58) conceives the institutional structure of science itself as tending to suppress new discoveries and encourage activities protecting vested intellectual interests. In the same vein, De Gre attributes the emergence of new knowledge to the "interaction of divergent social groups, each with its own dominant pattern of values..." Much of the evidence garnered by the sociology of knowledge tends to sustain the thesis that science is to some degree the product of the attitudes, interests, and values of the culture in which it exists. The discipline is relatively new and the available evidence inconclusive, but it would appear that the claims of the philosophers of science in regard to the university of the cognitive devices (to the extent that the term is interpreted the same) may need to be qualified in a temporal sense at least.

The initial motivations for inquiry are practical problems encountered in experience. A discipline is viewed as an area of investigation identified in terms of the objects it deals with and the point of view from which they are observed. Development of a discipline is a process of successive definition and successive integration of the objects and phenomena of experience. A discipline is further characterized as an area of investigation in which an approximate relationship between a limited number of objects and phenomena has been established. The primary objective of a discipline is the formulation of more precise statements of relationships. Since there is a reciprocal relationship between the process of abstracting (i.e. formulating statements of relationship) and observation of concrete entities, the literature of a discipline tends to consist of reports of the objects and phenomena being observed and abstractions or generalizations derived from meaningful observations. A peripheral aspect of the literature deals with the methods employed in making and interpreting observations. Although much of the foregoing discussion describes processes in the natural and social science disciplines, the humanities are related in the sense that they are concerned with the interpretation of artistic creations but not in establishing universal principles. The primary objective in instruction is to reconstruct the history and current status of a discipline. The usual method is to select "landmark" figures and contributions to indicate the problems with which the discipline is concerned, the observations which were made, and the interpretation given to these observations. Generally the student is given sufficient exposure through reading, lecture, and discussion to gain a feeling for the intellectual problems dealt with, the nature of the evidence employed, the techniques used in observation, and the reasoning processes applied to the data. The object would seem to be to equip the student to gain competence in collecting and interpreting further evidence as a practitioner in the field or at best to appreciate the tenuousness of factual information and the limitations in methodology in order that he may make appropriate judgments in the future.

If this is a correct interpretation of the instructional process, it may be observed that in synthesizing the literature of a discipline, the instructor extracts portions from the entire body of literature. The amount extracted is that, which in his judgment, is sufficient to explain a concept, illustrate the nature of the evidence and the method employed to establish it. He may emphasize the historical approach or concentrate on current problems, and he may use hypothetical data more meaningful to the students. The portions of the literature used may be highly variable from one instructor to another. He may vary his approach for the sake of novelty or some other reason, but he must not fail to convey the essence and the flavor of the discipline if he is to fulfill his mission. There are a variety of instructional devices employed in the transmission of information in college teaching. Some are books and articles written for that purpose, others are extracted from the subject literature, and others are created by the instructor. There would appear to be a core of "classics" or landmarks that every student must understand to attain proficiency in the "language" of a discipline.

In spite of the wide variation in stated objectives of colleges and the differences in style and methods of instructors, a number of factors tend to impose concensus in the content of the curriculum. The majority of faculty are trained in the major universities of the United States. Each faculty member is trained as specialist in his discipline. The discipline represents a body of information, regarded by concensus of the members of the discipline, as significant. In addition there are problem areas, reflecting its current status, which have not been resolved and on which there are varying opinions. A faculty member will endeavor to transmit to his students as much of the information about the discipline as he is able. His primary loyalty is to his discipline. He recognizes that his standing in the discipline is dependent upon the leaders in the discipline.

A college may take a variety of routes in teaching a discipline, but will be limited, in most cases, by the standards expected by graduate school. A college values the acceptance of its students for graduate work and gears its procedures accordingly. A further measure of a college's success is the acceptance of its graduates by employers and practitioners. All of these factors tend to impose standards of performance on college instruction. Accrediting agencies, as well, enforce at least a minimum level of quality on programs.

The methods of acquiring and transmitting information in college instruction suggests that there is a core of literature common to each subject. Further investigation into the structure of the literature of a discipline may be fruitful in identifying its characteristics and extent.

Appendix II

Citation Analysis of "Social Stratification"

One way to establish the existence of a core is to exhibit it. The inability to exhibit something does not prove its non-existence, but the ability to exhibit it establishes its validty. Consequently, an effort was undertaken to demonstrate the existence of a core by exhibiting at least a part of it. A subsidiary hypothesis was proposed and tested, namely that the citations of the author of textbooks and monographs in a given discipline would reveal substantial agreement about the literature which was essential to an understanding of that discipline. This hypothesis was tested on a small scale in a particular subject, social stratification, which forms part of the study of sociology. The field of sociology was chosen as being a discipline having some of the characteristics of both the sciences and the humanities, and because it is widely represented in college curricula.

Two professors of sociology at Washington and Lee University were given the following instructions:

(1) Choose a particular subject area in the field of sociology which forms the basis of one or two chapters in most college texts written for the introductory course in sociology and, at the same time, forms the basis of a complete semester-length course at the advanced undergraduate level.

(2) Supply the names and authors of four texts written for the introductory course in sociology at the college level, and which devote one or two chapters to the subject chosen.

(3) Supply the names and authors of four texts written for an advanced undergraduate course on the subject chosen.

(4) Supply the names of six to eight monographs which would be suitable supplementary or parallel reading in the above advanced course.

Both professors responded with the required lists, one on the subject of social stratification, the other on Population Analysis. Due to limitation of time, the analysis that is described below was applied to the subject of social stratification only. The books on this list were loaned by the professor or secured from the library. Each of these books was assigned a code consisting of a Roman numeral, followed by a capital letter. The Roman numerals were used to indicate: (I) an introductory level text, (II) an advanced text, and (III) a monograph. The capital letter served to distinguish the different texts within each category. The sixteen books and their code designations are shown in Table VIII.

Table VIII

Introductory and Advanced Texts and Monographs on Social Stratification in Citation Analysis

Code
Introductory Texts:

I-A Lunberg, Schrag, Larsen and Catton, *Sociology* (4th edition) New York: Harper and Row, 1968.
I-B Chinoy, *Society: An Introduction To Sociology* (2nd edition)
I-C Smelser (ed) *Sociology: An Introduction* New York: John Wiley and Sons, Inc., 1967.
I-D Broom and Selznick, *Sociology: A Text With Adapted Readings* (4th edition) New York: Harper and Row, 1968.
Advanced Texts:
II-A Hodges, *Social Stratification: Class In America* Cambridge: Schenkman Pub. Co. 1967.
II-B Kahl, *The American Class Structure* New York: Holt, Rinehart and Winston, 1966.
II-C Barber, *Social Stratification: A Comparative Analysis of Structure and Process* New York: Harcourt, Brace, and Co., 1957.
II-D Reissman, *Class in American Society*. Glencoe: The Free Press, 1959.
Monographs:
III-A Hollingshead, *Elmtown's Youth* New York: John Wiley and Sons (science edition) 1961.
III-B Bendix and Lipset, *Social Mobility In Industrial Society*. Berkeley: Univ. of Calif. Press, 1964.
III-C Blau and Duncan, *The American Occupational Structure* New York: John Wiley and Sons, 1967.
III-D Baltzell, *The Protestant Establishment* New York: Random House (Vintage Books) 1966.
III-E Dahrendorf, *Class and Class Conflict In Industrial Society* Stnaford: Stanford University Press, 1959.

III-F Mills, The Power Elite New York: Oxford Univ. Press (Galaxy Books) 1959.

III-G Vidich and Bensman, Small Town In Mass Society Garden City: Doubleday Anchor Books, 1960.

III-H Lenski, Power and Privilege, New York: McGraw-Hill, 1966.

These books were then examined for citations. Every citation in the four advanced texts, in the eight monographs, and in the appropriate chapters of the four introductory texts were copied separately onto 3 x 5 file cards, and assigned a code consisting of the code designation of the text from which the citation was obtained, plus an arabic numeral. Thus the code II-B-117 identifies the 117th citation in the advanced text by Kahl: The American Class Structure. The total number of citations obtained in this way was 3089. The citations from each source were then examined for duplications. Where a single author cited the same book or article more than once, the duplicate citations were then removed. This process eliminated 161 cards.

These remaining 2928 were then alphabetized by author. In this way duplicate citations by different authors were juxtaposed. The number of distinct citations by all sixteen authors was 2325. In the analysis to follow, this is taken to be the "universe" of books and articles on the subject of social stratification. Where two or more of the primary authors cited the same work, the appropriate code numbers of the set of two, three, four, etc., cards were recorded. Of the 2325 distinct citations, 393 were cited by two or more of the primary authors, and 1932 were cited by only one author. The first ten titles of the working chart of the 393 duplicate titles is shown in Table IX.

Table IX

Sample List of Duplications in Citation Analysis

Duplication No.	Duplicate Citations
1.	II-C-1 and II-B-120
2.	II-C-3 and III-B-245
3.	III-B-454 and II-A-2
4.	III-H-7 and II-A-3
5.	III-F-86 and II-B-40
6.	II-B-100 and II-A-5

7. II-A-6 and II-C-6
8. II-D-178, II-C-7 and II-A-7
9. III-D-56 and III-F-90
10. III-B-188 and II-B-85

Only one book was cited by as many as 7 of the 16 authors. Table X gives the breakdown for the other categories.

Table X

Frequency of Citations in Sociology Sample

No. Times Title Cited	No. Titles Cited
2	254
3	89
4	34
5	12
6	3
7	1

It is appropriate to remark here that different editions of a book and different pagination from the same book were treated as duplicate citations. Judgment was called for in a number of other situations; the general philosophy in making a judgment was to err on the side of assigning two different versions as being duplicates.

Next a matrix of duplicate citations was prepared as shown in Table XI. The number at the intersection of any row and column is the number of citations in common by the two primary authors having the code numbers indicated. As an example, 66 out of the 457 works cited by Barber (II-C) in his advanced text: *Social Stratification: A Comparative Analysis of Structure and Process*, were also among the 358 citations in the monograph by Lipset and Bendix: *Social Mobility in Industrial Society* (III-B). Also shown in each intersecting square is a number in parenthesis. This is the number of duplicate citations to be expected between the two authors if their citations had been selected from the 2325 titles of the "universe" at random. These numbers have been rounded to the nearest whole number. The meaning of these numbers can be made more graphic by the following description of the kind

Table XI - Complete Matrix of Duplicate Citations

	I				II				III							
	A	B	C	D	A	B	C	D	A	B	C	D	E	F	G	H
I-A	39	0 (1)	0 (0)	0 (1)	4 (7)	1 (2)	1 (8)	0 (3)	0 (1)	1 (6)	0 (2)	0 (2)	0 (3)	1 (5)	0 (0)	3 (8)
I-B	0 (1)	35	0 (0)	0 (1)	4 (6)	1 (2)	3, (7)	1 (3)	0 (0)	1 (5)	1 (2)	1 (2)	2 (3)	2 (4)	0 (0)	4 (7)
I-C	0 (0)	0 (0)	26	1 (1)	3 (5)	1 (1)	2 (5)	1 (2)	0 (0)	0 (4)	2 (1)	0 (2)	0 (2)	0 (3)	0 (0)	4 (5)
I-D	0 (1)	0 (1)	1 (1)	54	11 (10)	2 (3)	15 (11)	13 (4)	2 (1)	9 (8)	4 (3)	0 (3)	1 (5)	0 (6)	0 (0)	12 (11)
II-A	4 (7)	4 (6)	3 (5)	11 (10)	411	19 (22)	95 (81)	52 (34)	5 (5)	44 (63)	8 (21)	7 (26)	17 (36)	7 (47)	1 (1)	30 (82)
II-B	1 (2)	1 (2)	1 (1)	2 (3)	19 (22)	124	27 (24)	12 (10)	3 (2)	11 (19)	3 (6)	2 (8)	8 (11)	6 (14)	0 (0)	12 (25)
II-C	1 (8)	3 (7)	2 (5)	15 (11)	95 (81)	27 (24)	457	73 (38)	7 (6)	66 (70)	10 (23)	3 (29)	24 (40)	11 (53)	0 (1)	33 (91)
II-D	0 (3)	1 (3)	1 (2)	13 (4)	52 (34)	12 (10)	73 (38)	191	4 (3)	34 (29)	5 (10)	1 (12)	12 (17)	5 (22)	1 (0)	13 (38)
III-A	0 (1)	0 (0)	0 (0)	2 (1)	5 (5)	3 (2)	7 (6)	4 (3)	31	3 (5)	0 (2)	0 (2)	1 (3)	1 (4)	0 (0)	1 (6)
III-B	1 (6)	1 (5)	0 (4)	9 (8)	44 (63)	11 (19)	66 (70)	34 (29)	3 (5)	358	11 (18)	3 (22)	15 (32)	7 (41)	0 (0)	21 (71)
III-C	0 (2)	1 (2)	2 (1)	4 (3)	8 (21)	3 (6)	10 (23)	5 (10)	0 (2)	11 (18)	118	0 (7)	2 (10)	0 (14)	0 (0)	12 (24)
III-D	0 (2)	1 (2)	0 (2)	0 (3)	7 (26)	2 (8)	3 (29)	1 (12)	0 (2)	3 (22)	0 (7)	145	3 (13)	5 (17)	0 (0)	3 (29)
III-E	0 (3)	2 (3)	0 (2)	1 (5)	17 (36)	8 (11)	24 (40)	12 (17)	1 (3)	15 (32)	2 (10)	3 (13)	205	6 (24)	0 (0)	16 (41)
III-F	1 (5)	2 (4)	0 (3)	0 (6)	7 (47)	6 (14)	11 (53)	5 (22)	1 (4)	7 (41)	0 (14)	5 (17)	6 (24)	268	2 (0)	12 (53)
III-G	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)	0 (1)	1 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (0)	3	2 (1)
III-H	3 (8)	4 (7)	4 (5)	12 (11)	30 (82)	12 (25)	33 (91)	13 (38)	1 (6)	21 (71)	12 (24)	3 (29)	16 (41)	12 (53)	2 (1)	463

of process that could be carried out to obtain them. Suppose the universe deck of cards containing the 2325 titles of works cited by the sixteen authors were first given to the author of the text I-A, and he were asked to choose 39 at random from the deck. (39 is the actual number of distinct citations by the author of text I-A). After recording the names of the 39 titles as chosen, these 39 cards are returned to the deck. Assume the deck is then shuffled and presented to the author of the advanced text II-A, and he were asked to choose 411 cards at random, 411 being the actual number of citations by this author in his text. On the basis of probability theory, how many of these 411 should one expect to be on the list of 39 chosen earlier? To a very good approximation, the answer is (39 x 411)/2325, or 6.9 which, rounded off to the nearest whole number, is 7. The actual number of citations in common between the two authors was 4. It will be seen that in all but a few cases the actual number of duplicate citations was less than the number to be expected on the basis of random selection.

There are at least two reasons for caution in interpreting the results. The first is that the figures given above take no account of the publication dates of the citing and the cited works. It is obvious that an author cannot cite a work published subsequent to his own or, indeed for possibly a year or two prior to the publication date of his work. The effect of this limitation would be to reduce the number of duplicate citations to be expected on the basis of random selection. An exhaustive analysis of this effect was not undertaken. The publication dates of the primary sources were noted and the date of the most recent citation by each of the sixteen authors was also identified. A few cases were scanned with the result that the reductions to be expected from an exact analysis were judged not to significantly affect the general conclusion that there is no concensus among these sixteen authors that would justify reliance upon citation analysis as a means of identifying a core collection.

A second effect was also examined superficially. The "universe" of 2325 citations includes books, journal articles, articles from popular magazines, U.S. statistical tables, etc. It might be that if one confined his attention to duplications of books only, the ratio of actual to probably duplications would be higher. A rough count of the "universe" deck revealed that about two-thirds of the citations were books. If duplicate citations of works other than books were removed, the "universe" is smaller, but the number of duplicate citations is also smaller. Again, limiting the analysis to books only did not promise to affect the overall results significantly.

Several disparate conclusions might be inferred from the result of this study. First, it may suggest that the core collection is a small fraction of the total literature, comparable perhaps to the number that would be obtained from random selection. It would be appropriate to obtain expert judgment as to whether the duplicate citations actually obtained constitute a suitable "core" or material in this subject. If this conclusion is valid, then on the basis of this study, the "core" size would be about 15% of the "universe" of literature on that subject.

The study might be regarded as revealing more about the citing habits of authors than of the literature of a discipline. For example, in only two cases did as many as two of the authors of elementary texts agree on a citation, whereas in 201 cases one or more authors of elementary texts were in agreement with authors of the eight monographs. It is likely that authors cite for different reasons, among them being the possibility that his citation should be distinctive rather than utilitarian. However, the authors of the four advanced texts agreed on 176 titles in two or more cases.

In summary, this particular study failed to achieve the desired result--namely to exhibit the core of literature in a particular subject. At worst, it demonstrates the futility of determining a core by citation analysis. At best, it raises fundamental questions about the utility of citations, particularly at the elementary text level, and suggests several areas for fruitful investigation. Perhaps, too, it will suggest to some readers alternative approaches for determining a core collection. The null result of this experiment is the basis for the conclusion, given in the body of this report, that the most feasible method for arriving at a core collection is through a concensus of subject matter experts.

References and Bibliography

1. *A Feasibility Study for a Joint Computer Center for Five Washington, D. C. University Libraries. Final Report.* Consortium of Universities of Metropolitan Washington, D. C. 1968.

2. Arthur D. Little, Inc. *A Plan for a Library Processing Center for the State University of New York.* Report to the Office of Educational Communications, State University of New York. November, 1967.

3. Bell, Daniel, "The Disfunction of Culture and Social Structure," *Daedalus*, 94 (Winter, 1965).

4. Benjamin, Cornelius, *An Introduction to the Philosophy of Science* (New York: The Macmillan Co., 1937).

5. Bernal, John Desmond, *Science in History* (London: Watts 1954).

6. Berry, Edwin R., *Centralized Processing for Systems of Libraries*, Albany, N. Y. New York State Library n.d.

7. Borko, Harold, *Automated Language Processing* (New York: Wiley, 1967).

8. Bourne, Charles P., *Methods of Information Handling* (New York: Wiley, 1966).

9. Bridgman, P. W., *The Logic of Modern Physics* (New York: The Macmillan Co., 1927).

10. Brubacher, John S. and Rudy, Willis, *Higher Education in Transition* (New York: Harper and Brothers, c1958).

11. Brutcher, Constance, Gassford, G., and Rexford, E., "Cost Accounting for the Library," *Library Resources and Technical Services*. 8:413, Fall 1964.

12. California State Colleges. *Second Report to the Chancellor from the Committee on Library Development for the California State Colleges*. 1966.

13. Carter, Launor F., *et. al.*, *National Document - Handling Systems for Science and Technology* (New York: Wiley, 1967).

14. Clapp, Verner W., "The Future of the Research Library" (Winsor Lecture, University of Illinois, 1963).

15. Clapp, Verner W., and Jordan, Robert T., "Re-evaluation of Microfilm As a Method of Book Storage," College and Research Libraries, 24, No. 1 (January, 1963).

16. College and Research Libraries, 29:330. July, 1968.

17. Colorado Academic Libraries Book Processing Center Study Final Report. Phase I and Phase II. Boulder, Colorado, University of Colorado, 1968.

18. Conant, James B., Science and Common Sense (New Haven: Yale University Press, 1951).

19. Cooper, Marianne, "Criteria for Weeding of Collections," Library Resources and Technical Services, 12:3, 339-351. Fall, 1968.

20. Cuadra, Carlos A., ed. Annual Review of Information Science and Technology (New York: Wiley, 1966) I.

21. Darling, Richard L., "School Library Processing Centers," Library Trends, 16:60. July, 1967.

22. Development of a Computer Processing Center for the New England State University Libraries, Submitted to the New England Board of Higher Education and the Council on Library Resources. Inforonics, Inc. 1967.

23. Dix, William S., "New Challenges to University Libraries," University: A Princeton Quarterly, No. 26 (Fall, 1965).

24. Downs, Robert B., A Survey of Cooperating Libraries for the Kansas City Regional Council for Higher Education. Kansas City, Regional Council for Higher Education, 1964.

25. Educational Facilities Laboratories, The Impact of Technology on the Library Building (New York: Educational Facilities Laboratories, 1967).

26. Education at Berkeley: Report of the Select Committee on Education (Berkeley, California: University of California, Academic Senate, March 1966).

27. "Guidelines for Centralized Technical Services." Library Resources and Technical Services, 10:233, Spring, 1966.

28. Hempel, Carl G., "Fundamentals of Concept Formation in Empirical Science," International Encyclopedia of Unified Science, II, No. 7 (Chicago: The University of Chicago Press, 1952).

29. Henderson, Algo D., Policies and Practices in Higher Education (New York: Harper and Brothers, c1960).

30. Hiatt, Peter, "Cooperative Processing Centers for Public Libraries," Library Trends, 16:73, July, 1967.

31. Hunt, James R., "The Historical Development of Processing Centers in the United States," Library Resources and Technical Services, 8:54, Winter, 1964.

32. "The Information Deluge," The Johns Hopkins Magazine (Fall, 1967).

33. Lazarsfeld, Paul F., and Rosenberg, Morris, The Language of Social Research (Glencoe, Illinois, Free Press, 1955).

34. Leimkuhler, Ferdinand F., and Neville, Anthony E., "The Uncertain Future of the Library," The Johns Hopkins Magazine (Fall, 1967).

35. Lenzen, Victor, "Philosophy of Science," In Dagobert D. Runes, Twentieth Century Philosophy (New York: Philosophical Library, 1943).

36. Lenzen, Victor, "Procedures of Empirical Science," International Encyclopedia of Unified Science, I, No.5 (Chicago: The University of Chicago Press, 1938).

37. Leonard, Lawrence E., Cooperative and Centralized Cataloging and Processing: A Bibliography. University of Illinois Graduate School of Library Science. Occasional Papers No. 93. July, 1968.

38. Lundberg, George A., Foundations of Sociology (New York: The Macmillan Co., 1939).

39. McCune, Lois C., and Solomon, Stephen R., "Bibliography of Library Automation," ALA Bulletin, 61, No. 6 (June, 1967).

40. McQuarrie, Catherine, "Cost Survey; Cost of Ordering, Cataloging and Preparations in Southern California Libraries." *Library Resources and Technical Services*, 6:337, Fall, 1960.

41. Merton, Robert K., *Social Theory and Social Structure* (Glencoe, Illinois: Free Press, 1949).

42. Merton, Robert K., "Sociology of Knowledge," In George Gurvitch and Wilbert E. Moore, ed. *Twentieth Century Sociology* (New York: Philosophical Library, 1945).

43. Nelson Associates, Inc. *Feasibility of School and College Library Processing through Public Library Systems in N. Y. State. A Report for the New York State Library*. 1966.

44. Northrop, F.S.C., *The Logic of the Sciences and the Humanities* (New York: Macmillan Co., 1947).

45. Parsons, Talcott, *Structure of Social Action* (Glencoe, Illinois: Free Press, 1949).

46. Patterson, Franklin, and Longsworth, Charles R., *The Making of a College* (Cambridge, M.I.T. Press, 1966).

47. Polanyi, Michael, *Science, Faith and Society* (London: Oxford University Press, 1946).

48. Rider, Fremont, *The Scholar and the Future of the Research Library* (New York: Hadham Press, 1944).

49. Rudolph, Frederick, *The American College and University* (New York: Knopf, 1965).

50. Shores, Louis, "The Liberal Arts College, A Possibility in 1954?" *School and Society*, Vol. XLI (January 26, 1935).

51. Toffler, Alvin, "Libraries," In Educational Facilities Laboratories, *Bricks and Mortorboards* (New York: Educational Facilities Laboratories, 1966).

52. Vann, Sarah K., Southeastern Pennsylvania Processing Center. Pennsylvania State Library Monograph No.4, 1967.

53. Verry, H.R., and Wright, G. H., Microcopying Methods (Rev. ed., London: Focal Press, 1967).

54. Vickery, B. C., On Retrieval System Theory (2nd ed.; London: Archon Books, 1965).

55. Werkmeister, W. H., The Basis and Structure of Knowledge (New York: Harper and Brothers, 1948).

56. Wynar, Donald, "Cost Analysis in a Technical Services Division." Library Resources and Technical Services. 7:312. Fall, 1963.

57. Zetterberg, Hans Lennart, On Theory and Verification in Sociology (New York: Tressler Press, 1954).

58. Znaniecki, Florian, The Method of Sociology (New York: Farrar & Rinehart, 1934).